THE
PHYSICAL REVIEW
A Journal of Experimental and Theoretical Physics Established by E. L. Nichols in 1893
VOL. 57, NO. 2 JANUARY 15, 1940 SECOND SERIES

The Production of Soft Secondaries by Mesotrons

J. R. OPPENHEIMER AND H. SNYDER, *University of California, Berkeley, California*
AND
R. SERBER, *University of Illinois, Urbana, Illinois*
(Received November 3, 1939)

I. Cosmic-ray evidence shows that the soft component, which persists even under great thicknesses of matter, consists, primarily at least, of electrons and γ-rays. The observations of the magnitude and variation with depth of the soft component and showers, and of the size and material dependence of bursts, lead to the following rough phenomenological description: In addition to ordinary ionization losses, the mesotrons have an appreciable chance of transferring a considerable fraction of their energy to the soft component. For transfers above 2×10^{10} v this probability is roughly independent of mesotron energy. Under 10^{10} v the probability of large transfers is much greater, roughly 20 times as great as at higher energies. II. The production of secondaries of energy $<10^{10}$ v can be accounted for by a familiar process: the Lorentz contracted Coulomb field of the mesotron, in sweeping over an atom, ejects a high energy electron. This mechanism, however, is inadequate to explain the very large energy transfers, $>10^{10}$ v, involved in bursts. The theory which has been developed to describe the mesotron of spin one associates with transitions in which the direction of the mesotron spin changes an intrinsic dipole moment; this dipole field favors high energy transfers, and, roughly, can account for the observed bursts. However, the dipole field also gives probabilities of high energy bremsstrahlung, $>10^{10}$ v, much too large to be compatible with the observations. III. The cross sections for the production of electron secondaries and bremsstrahlung by the intrinsic mesotron dipole field have been estimated under the assumption that the coupling of mesotron and electromagnetic field is small; when this is not so the formula cannot be right. Examination of the approximation made indicates that the estimate of the cross section for electron secondaries should be valid up to mesotron energies of $\sim10^{12}$ v, and is thus applicable to the bursts. The bremsstrahlung formula, on the other hand, fails at 10^{10} v; it thus cannot be used for bursts, while for energies $10^9<E<10^{10}$ it leads to no contradiction with the experimental evidence. The problem of extending the formulae above these critical energies probably goes beyond the framework of present theory. The evidence indicates that it is the largeness of the coupling, and not the occurrence of lengths smaller than the critical $\hbar/\mu c\sim2\times10^{-13}$ cm, that limits the applicability of the quantum mechanics.

I. INTRODUCTION: OBSERVATIONS

THE persistence of the soft component of cosmic radiation under great thicknesses of matter shows that this is secondary radiation produced by the mesotrons of the penetrating component. At sea level the cascade radiation which comprises the soft component still has three sources of roughly equal importance:[1] degraded primary cascade radiation of high initial energy; secondaries produced by the mesotrons in matter; and (less certainly) secondaries from the disintegration of mesotrons in the atmosphere. But after a few meters water equivalent below sea level only the second group of secondaries is of importance.

From this point on, the fraction of soft com-

[1] We are grateful to Professor E. Amaldi for informing us of unpublished experimental results. For a general discussion, see Euler and Heisenberg, Ergeb. d. exakt. Naturwiss. **17**, 1 (1938).

ponent increases with depth.[2] Of the order of 10 percent of the ionizing particles are soft at 10 m under sea level; the proportion is about doubled at 30 m. The number of showers and of bursts also tails off less rapidly with filtration than the number of mesotrons. At 75 m below sea level the relative number of showers is about four times that at 10 m. This soft radiation has the typical behavior of cascade radiation of electrons and γ-rays, and there is no evidence that heavy particles or low energy mesotrons play an important part in it.

These results may be simply interpreted: filtration increases the mean mesotron energy, and for a power law energy distribution this increase is linear with depth. Thus the increase in the soft component with filtration means that the mean energy transfer from mesotrons to soft component must increase with mesotron energy.

Some information on the energy transfer cross section in the region of very high energies is given by what is known of the variation of burst frequency with burst size and with material. A most striking result here[3] is that the distribution of bursts in size, above roughly 10^{10} v, follows the same law as the energy distribution of the penetrating mesotrons: the number of bursts greater than S in size is given

$$N_{>S}\sim S^{-\gamma}, \quad \gamma\sim 1.8, \tag{1}$$

by whereas the number of mesotrons of energy greater than E is given by the law, obtained from the absorption curve on the assumption that most of the absorption comes from ionization losses,

$$N_{>E}\sim E^{-\gamma}, \quad \gamma\sim 1.9. \tag{2}$$

This parallelism of the burst frequency and absorption curves follows immediately if for high energies the probability of a given fractional transfer of energy from a mesotron to soft radiation is independent of mesotron energy. Such a law of energy transfer does, in fact, give a mean energy loss just proportional to the mesotron energy.

Additional evidence on the form of the cross section is provided by observations on the relative burst frequencies in different materials. As is discussed in more detail elsewhere,[4] a cross section depending only on fractional energy loss gives, for the frequency of bursts in substances of different atomic number, a Z dependence

$$(g(Z)/Z)I(Z)^{1-\gamma}, \tag{3}$$

where $g(Z)$ gives the atomic number dependence per atom of the large energy transfers, and $I(Z)$ is the critical energy of the cascade theory, which increases with decreasing Z. It seems probable that, by taking into account the effect of scattering on the low energy cascade radiation, this result, if $g(Z)=Z$, can be reconciled with the observed approximate equality of burst frequencies in different materials.

These considerations suggest that, in addition to the ordinary ionization losses, mesotrons can transfer to the soft component an appreciable fraction of their total energy, and that the probability of this is roughly independent of the mesotron energy. It is, however, not possible to fix the constant determining this probability to give agreement both with the increase in soft radiation on filtration and with the absolute value for the probability that a high energy mesotron make a burst. To see this, let us write for the cross section per atom for a mesotron of energy E and mass μ to transfer an energy fE to the soft component

$$d\sigma=\sigma_0\kappa(f)df, \quad \sigma_0=\pi Ze^4/\mu^2c^4. \tag{4}$$

Then the probability that in a centimeter of material a mesotron of energy greater than E makes a burst of energy greater than E is just

$$\xi=\sigma_0 N\int_0^1 \kappa(f)f^\gamma df, \tag{5}$$

where N is the number of atoms per cc of the material. According to Schein and Gill[5] the probability that a mesotron will make a burst of energy greater than 2×10^{10} v is about $\xi=2\times10^{-5}$ per cm.

An independent check on this figure we may obtain by the following argument: the maximum of the transition curve for bursts in Pb, at $\sim$5 cm, is about twice the value at greater thick-

[2] W. M. Nielsen and K. Z. Morgan, Phys. Rev. **54**, 245 (1938); P. Auger and T. Grivet, Rev. Mod. Phys. **11**, 232 (1939).

[3] A. Sittkus, Zeits. f. Physik **112**, 626 (1939); M. Schein and P. S. Gill, Phys. Rev. **55**, 1111 (1939).

[4] J. R. Oppenheimer, Rev. Mod. Phys. **11**, 264 (1939).

[5] Schein and Gill, reference 3.

nesses. Thus the bursts produced by cascade radiation give a contribution equal to that from the mesotrons: decay electrons will not give an appreciable contribution. The total number of cascade electrons is about 1/20 the number of mesotrons. Of these the theory of showers shows that $(4\times10^7/\epsilon)^{1.9}$ have an energy greater than ϵ. On the other hand, if $\epsilon>10^{10}$, roughly $(2\times10^9/\epsilon)^{1.9}$ mesotrons have an energy greater than ϵ. The bursts they make have a "range" of about 5 cm in Pb. Thus we must have, for the probability ξ that a mesotron of energy $>\epsilon$ make a burst in a cm of Pb,

$$5\xi(2\times10^9/\epsilon)^{1.9}=(1/20)(4\times10^7/\epsilon)^{1.9},$$
$$\xi=(1/100)(1/50)^{1.9}\sim0.5\times10^{-5}.$$

The value $\xi\sim10^{-5}$, gives

$$\kappa=\int_0^1\kappa(f)f^2df\sim\tfrac{1}{3}. \qquad (6)$$

Now this value is not great enough to account for the increase in relative soft radiation with filtration. To see this we have to compare the total energy lost to the soft component per cm by (4) with the mean ionization energy loss of the secondaries, which gives then directly the number of soft particles per mesotron. This ratio is

$$R=m\bar{E}\int_0^1\kappa(f)fdf/4\mu^2c^2\ln\,(I(Z)/ZRh). \qquad (7)$$

Here m is the electron mass and $\bar{E}$ the mean mesotron energy. This gives, for the increase in ratio of soft to hard component on filtration by 30 m water equivalent, if we use (6) to estimate κ,[6] only 0.5 percent, whereas the experimental value is more nearly 10 percent.

Since the mesotron energies and energy transfers involved in the increase of the soft component are primarily less than 10^{10} v, and the bursts involve energies greater than 2×10^{10} v, we are led to the following rough phenomenological description: In addition to ordinary ionization losses, the mesotrons have an appreciable chance of transferring a considerable fraction of their energy to the soft component. For transfers above 2×10^{10} v this probability is roughly independent of mesotron energy and is given by (4) and (6). Under 10^{10} v the probability of large transfers is much greater, and in this range it is roughly 20 times as great as at higher energies.

II. Electron Secondaries and Bremsstrahlung

The elementary processes by which mesotrons transfer energy to the soft component may conveniently be classified as nuclear or electromagnetic, according to whether the coupling of mesotrons with the heavy particles or with the electromagnetic field plays the primary part in them. Because of the large nuclear coupling and high mesotron mass it has usually been assumed that the nuclear processes would alone be important for high energy transfers. In fact, calculations of nuclear effects, such as mesotron absorption and scattering by nuclei, calculations based on the perturbation theoretic treatment of the coupling as small, lead to cross sections so large, for high energies, that they completely contradict the high penetrating power of the mesotrons. It has been emphasized especially by Heisenberg[7] that the prediction of these large cross sections rests on the essentially erroneous treatment of the interactions as small; despite several attempts no reliable estimate of them has been given, and this problem probably goes beyond the framework of present theory. Nor is it sure in what way such impacts will give appreciable amounts of soft radiation: the emission of γ-rays from nuclei excited by such impacts hardly seems a very plausible mechanism; the radiative capture of a mesotron by a nucleus, inverse to the mesotron photo-effect, seems more probable.

Under these circumstances we have thought it profitable to re-examine the electromagnetic effects a little more closely. For the problem of energy transfers to the soft component two types of collision are of first importance: elastic impacts with free electrons, and bremsstrahlung: to this latter corresponds, for the inverse problem of the creation of mesotrons by the soft component, pair production.

[6] We have supposed $\int^1\kappa(f)f^2df\sim\int_0^1\kappa(f)fdf$; for a distribution $\kappa(f)$ which greatly favored the emission of slow electrons this would no longer be right. For example $\kappa(f)\sim1/f^2$ would increase the fraction of soft component of energy $>E_{min}$ by a factor $\ln(\bar{E}/E_{min})$.

[7] W. Heisenberg, Zeits. f. Physik **113**, 61 (1939).

If mesotrons satisfied the Dirac equation, cross sections for these two processes could be obtained from the familiar calculations for electrons. For the cross section per atom that a mesotron of energy E gives a fraction f of its energy to an electron,

$$d\sigma=\frac{2\pi Ze^4}{mc^2E}\frac{df}{f^2},\quad f<1/\left(1+\frac{\mu^2c^2}{2mE}\right)$$
$$=2\sigma_0(\mu^2c^2/mE)df/f^2. \tag{8}$$

Similarly, for the cross section per atom that a γ-ray of energy fE be emitted

$$d\sigma=\frac{4\alpha Z\sigma_0}{\pi}\ln\left[\frac{2E(1-f)}{\mu c^2Z^{\frac{1}{3}}f}\right]\left(f+\frac{4}{3}\frac{(1-f)}{f}\right)df, \tag{9}$$
$$\alpha=e^2/\hbar c.$$

In this result atomic screening has been neglected, but the Coulomb field has been smoothed out for $r<Z^{\frac{1}{3}}\hbar/\mu c$; and the Born approximation has, of course, been used.

The cross section (8) gives a logarithmic increase of the soft radiation with filtration; the fraction of the electrons of energy greater than $E_{\min}$ at depth t is

$$R=\tfrac{1}{2}\ln\left(\frac{2m\epsilon^2}{\mu^2c^2E_{\min}}\right)\Big/\ln\left(\frac{I(Z)}{ZRh}\right)$$
$$\sim\frac{\ln 10(\epsilon/10^9)^2}{36-4\ln Z}\quad\text{for}\quad E_{\min}=10^7\text{ v}, \tag{10}$$

where ϵ is the ionization energy loss of the mesotron in penetrating to the depth t. This gives, at sea level, $R=13$ percent, at 30 m below sea level, with $Z=12$, $R=25$ percent, at 80 m $R=31$ percent, in rough agreement with the observed values, although perhaps somewhat low at 80 m. From (9) we obtain $R\sim3\times10^{-5}Z(\epsilon/10^9)$, a much smaller contribution than (10).

For showers of energies $\times10^9$ v $<E<10^{10}$ v, (8) gives, instead of (1), the law $N_{>S}\sim S^{-\gamma-1}$. Both the size dependence and absolute number of showers predicted by (8) agree well with the experimental results.*

However, (8) has the wrong energy dependence, (9) the wrong Z dependence, to account for the observations on bursts. But it is interesting to note that both (8) and (9) give for bursts $>2\times10^{10}$ v, a probability in rough agreement with observation; thus (9) gives

$$\kappa\sim3\alpha Z\sim2\text{ for Pb},$$

whereas the experimental value is $\kappa\sim\frac{1}{2}$. Also (8) gives, for all Z and bursts $>2\times10^{10}$, a burst frequency corresponding to $\kappa\sim\frac{1}{2}$.

There is, however, valid ground for doubting the applicability to the burst processes of (8) and (9). For Yukawa's theory requires an integral mesotron spin, and the spin dependence of nuclear forces shows that this spin must be one. The theory which has been developed to describe these particles[8] makes the mesotron electromagnetic current density depend on derivatives of the mesotron field, and associates, with transitions in which the direction of the mesotron spin changes, current distributions more singular than those of a Dirac electron. The Fourier components of current corresponding to a momentum $\hbar k>\mu c$ behave like k^1; for the Dirac electron, and for mesotron transitions not changing the spin, they behave like k^{-1}. For a classical point charge they behave like k^0, and for a point dipole like k; thus one may roughly ascribe these singular currents to an intrinsic mesotron dipole moment. The Fourier components of current corresponding to this dipole moment are larger by a factor $k/\mu c$ than those corresponding to the electric charge.

These singular currents of course interact very strongly with high frequency radiation fields. They radically alter the high energy cross sections, giving much larger values than (8) and (9). At the same time they introduce couplings so large that the question of the perturbation theoretic estimate of the cross section requires re-examination.

This behavior of the cross sections is well illustrated by the work of Laporte[9] on the Coulomb scattering of mesotrons. Treating this problem by the Born approximation, one finds for scattering without spin change the Rutherford result

$$d\sigma=\tfrac{1}{2}Z\sigma_0(\mu c^2/E)^2\csc^4\tfrac{1}{2}\theta d(\cos\theta),\quad E\gg\mu c^2. \tag{11}$$

* See A. C. B. Lovell, Proc. Roy. Soc. **172**, 583 (1939).

[8] Yukawa, Sakata, and Taketani, Proc. Phys.-Math. Soc. Japan **20**, 319 (1938); Yukawa, Sakata, Kobayasi, and Taketani, Proc. Phys.-Math. Soc. Japan **20**, 720 (1938); N. Kemmer, Proc. Roy. Soc. **A166**, 127 (1938); Fröhlich, Heitler, and Kemmer, Proc. Roy. Soc. **A166**, 154 (1938); H. J. Bhabha, Proc. Roy. Soc. **A166**, 501 (1938).

[9] O. Laporte, Phys. Rev. **54**, 905 (1938).

For scattering with spin change

$$d\sigma = \tfrac{1}{3} Z\sigma_0 \operatorname{ctg}^2 \tfrac{1}{2}\theta d(\cos\theta). \qquad (12)$$

For elastic impacts with electrons one finds,[10] again using the Born approximation, an analogous result: for no spin change we again get (8); for spin change

$$d\sigma = \tfrac{2}{3}\sigma_0(1 - f + \tfrac{1}{2}f^2)df/f, \quad f < 1/(1+\mu^2c^2/2mE). \qquad (13)$$

This result is indeed of the form (4), and thus can account roughly for the size and material-dependence of large bursts. It gives

$$\kappa = 0.22.$$

This is of the order of magnitude observed, and together with (8) can probably account for the greater part of the high energy bursts.

For the bremsstrahlung, the high probability of large angle scattering given by (12), and the large coupling with high frequency radiation fields, give a cross section increasing rapidly with energy. For an unscreened Coulomb field the cross section takes the asymptotic form $\sigma \sim \alpha Z\sigma_0(E/\mu c^2)^2$; most of these are close impacts and correspond to large angles of scattering and radiation. It is therefore essential to take into account the modification of the Coulomb field within the nuclear radius $\sim Z^{\frac{1}{3}}\hbar/\mu c$. If we do this by eliminating all Fourier components of the Coulomb field with $k > \mu c/\hbar Z^{\frac{1}{3}}$, and thus ignoring nuclear collisions, we get for high E, still using the Born approximation, a cross section of the form

$$\sigma \sim \alpha Z^{\frac{1}{3}}\sigma_0 E/\mu c^2. \qquad (14)$$

It is clear that neither the E nor Z dependence of this cross section is in agreement with the observations on bursts. For $E > 2\times10^{10}$ v, (14) cannot be right. However, in the range 10^9 v $< E < 10^{10}$ v, (14) is in no contradiction to the experimental evidence: its contribution to the soft component would be only $R \sim 1$ percent.

We see, so that (13) can account roughly for soft radiation with $E > 2\times10^{10}$ v; that (14) can surely not be right for $E > 2\times10^{10}$ v. On what grounds can we regard (14) as invalid; and do these apply, and if so in what range of energies, to (13)?

[10] This result has been independently derived by Corben and Massey, Proc. Phil. Camb. Soc. in press. We are indebted to Dr. Corben for telling us of his results.

III. Validity

Two distinct but related criteria[11] are involved in the question of the validity of both our results (13) and (14) for secondaries and bremsstrahlung: Formulae for mesotron charge and current density have been used for arbitrary small wavelengths; one may question, as has often been done, whether they are right for distances $\ll \hbar/\mu c$. On the other hand, we have throughout treated the coupling of mesotron and electromagnetic field as small, have used, that is, a Born approximation: under what conditions is this treatment justified? The formulae (13) and (14) will surely have no validity when it is not.

The former condition is not in Lorentz-invariant form, nor can it be reformulated to apply only to the rest system of the mesotron, since in the impact the mesotron momenta change. A natural extension of the criterion is that only for those impacts in which momentum transfer π and energy transfer ϵ satisfy

$$\pi^2 - \epsilon^2/c^2 \gg \mu^2c^2 \qquad (15)$$

can charge and current expressions of mesotron theory be used. This condition is thus clearly equivalent to the rejection of relativistic mesotron theory. It is incisive; there is no evidence that it is right.

To apply this condition we note that in the rest system of the center of mass of electron and mesotron the energy of the electron does not change, and the transverse and longitudinal momentum changes are of the same order of magnitude. Since the transverse momentum transfer is invariant, and is of the order $(2mE)^{\frac{1}{2}}$, we get from (15)

$$E \gg \mu^2c^2/2m \sim 10^{10} \text{ v}.$$

It is, as we have seen, only above this energy that the terms (13) become of importance compared to (8).

In the same way we may apply (15) to the bremsstrahlung. If the Coulomb field were not

[11] W. Heisenberg, Zeits. f. Physik **110**, 251 (1938).

cut off at the nuclear radius, the scattering and radiation would be at large angles in the nuclear rest system, and (15) would give

$$E^2 \gtrsim \mu^2 c^4.$$

For the screened field the scattering and radiation is roughly isotropic in a coordinate system S where the mesotron momentum is $(2\mu E)^{\frac{1}{2}}$, so that $\mu E \gtrsim \mu^2 c^2$; in either case the bremsstrahlung formulae would have no range of validity at all, since they are for the case $E \gg \mu c^2$, and (15) denies that mesotron theory is applicable to such a problem.

For nuclear problems, where the relative magnitude of the coupling energy compared to the kinetic energy of the mesotrons is measured by $gE/\mu c^2$, with g a constant of order unity, the smallness of the interaction energy essentially requires (15). For electromagnetic effects the relative magnitude of the coupling is $\alpha^{\frac{1}{2}}E/\mu c^2$, $\alpha = e^2/\hbar c$; and the coupling may remain small even when (15) is violated.

It is clear from (13) and (14) that for high enough energies the coupling terms are not small. If we use for each problem the coordinate systems in which scattering is roughly isotropic, in which therefore for secondaries and bremsstrahlung, respectively, the mesotron momentum P is $(2mE)^{\frac{1}{2}}$ and $(2\mu E)^{\frac{1}{2}}$, the validity of the Born approximation at least requires that the cross section be small compared to the summed areas of the partial waves involved. For the secondaries, with nearly spherical scattering, this means that $\sigma < (\hbar/P)^2$, and thus that

$$E < \mu^2 c^2/mZ\alpha^2 \sim 4\times 10^{14}/Z \text{ v}. \qquad (16)$$

For the bremsstrahlung, the half-width, $\delta\theta$, of the angular distribution of the scattered mesotron, for fixed direction of emission of the γ-ray, is of the order $(\delta\theta)^2 \sim (\mu c/Z^{\frac{1}{2}})^2/\mu E \sim \mu c^2/Z^{\frac{1}{2}}E$; thus partial waves up to $L^2 \sim Z^{\frac{1}{2}}E/\mu c^2$ are involved; we must have $\sigma < (\hbar L/P)^2$ and

$$E < \mu c^2/\alpha^2 Z \sim 3\times 10^{14}/Z \text{ v}. \qquad (17)$$

To see whether these conditions, which are certainly *necessary* for the smallness of the coupling, are also sufficient, let us look more closely at the coupling energy density and the kinetic energy density. This is simplest in the coordinate systems S, where one may use wave packets of approximate dimensions $\hbar/P$ throughout. The kinetic energy density is then of the order $(P/\hbar)^3 Pc$ whereas the coupling energy, for those transitions involving a change of mesotron spin, is $\sim (P/\hbar)^3 \alpha^{\frac{1}{2}} (P/\mu c) Pc$, and the ratio is of the order

$$\alpha^{\frac{1}{2}} P/\mu c. \qquad (18)$$

Then we must have

$E < \mu^2 c^2/m\alpha \sim 2\times 10^{12}$ v for the validity of (13),

and

$E < \mu c^2/\alpha \sim 10^{10}$ v for the validity of (14).

These conditions are more incisive than (16,17), because the coupling of the mesotron with its radiation field becomes large before the coupling with the weaker fields of electron and nucleus do. Thus in the problem of the bremsstrahlung the maximum value of the electric field of the nucleus (in coordinate system S) acting on the mesotron is $Ze(Z^{\frac{1}{2}}\hbar/\mu c)^{-2}(E\mu/\mu c)^{\frac{1}{2}} = Z^{\frac{1}{2}}eE^{\frac{1}{2}}\mu^{\frac{1}{2}}c/\hbar^2$, whereas the zero point electric field fluctuations of wave-length $\hbar/P$, which induce mesotron radiation of this wave-length, are of the order $(\hbar c)^{\frac{1}{2}}P^2/\hbar^2 = (\hbar c)^{\frac{1}{2}}\mu E/\hbar^2$, and thus $(E/\mu c^2 \alpha Z^{\frac{1}{2}})^{\frac{1}{2}}$ larger.

The derivation of the results (13) and (14) can certainly then not be justified for energies greater than 10^{10} v for the bremsstrahlung, 2×10^{12} v for the secondaries. For to extend these results to higher energies would involve *not* treating the coupling of mesotrons and radiation as small. It is well known that this program leads to divergent results, and that only the roughest correspondence-theoretic analogies offer any guide to their interpretation. It should be emphasized that it would be of no use to treat more strictly, say in the problem of bremsstrahlung, the motion of the mesotron in the Coulomb field; it is not primarily the magnitude of *this* coupling which makes the trouble.

This discussion therefore suggests that the bremsstrahlung formula (14) may well be right up to 10^{10} v; at much higher energies the observational material on burst frequency shows that it must be wrong. On the other hand (13), for high energy secondaries, can be justified up to 2×10^{12} v, and, as we have seen, can explain at least a good part of the bursts observed.

It may be mentioned that for the inverse problem of mesotron production, pair production in a "cut-off" Coulomb field again gives a cross section of the form $\sigma=\alpha Z^2\sigma_0 E/\mu c^2$, and again a limit $\sim 10^{10}$ v for its validity. At this highest energy the cross section per atom of oxygen is 10^{-28} cm^2, which is about 300 times too small to account for the observed high intensity of secondary mesotrons of this energy. A cross section 300 times larger would, as a matter of fact, be inconsistent with the high penetrability of mesotrons. The large observed discrepancy between mesotron production and absorption probabilities has been repeatedly emphasized.[12]

[12] See for example, L. W. Nordheim, Phys. Rev. **56**, 502 (1939).

JANUARY 15, 1940 PHYSICAL REVIEW VOLUME 57

The Proton-Deuteron Transformation As a Source of Energy in Dense Stars*

W. A. WILDHACK
George Washington University, Washington, D. C.

(Received November 4, 1938)

The rates of energy evolution due to the transformation to helium, starting with the reaction $H+H=D+e^+$, in hydrogen at densities of 10^4 to 10^8 g/cm^3, were calculated on the basis of complete degeneracy, and the assumption of a crystal-like spacing of the protons. The results indicate that any considerable amount of hydrogen in white dwarf stars would lead to much higher luminosities than those observed. Thus the low effective molecular weight (1.5) as calculated for some of these stars from the accepted white dwarf model, cannot be due to a high content of hydrogen. It might be explained as due to very large content (~100 percent) of the helium isotope He^3 but it is very difficult to see how such large amounts of this isotope could be present in these stars. It appears that the paradox can be removed only by revision of the observational data concerning the white dwarf radii.

1. INTRODUCTION

THE importance of the reaction $H+H=D+e^+$ as a major source of stellar energy has been established by Bethe and Critchfield,[1] at least for stars lighter than the sun. In view of prevailing theories of stellar composition, which indicate a considerable proportion of hydrogen in most stars, it is desirable to investigate the rate of energy production due to this H−H reaction in dense stars.

Since the zero-point energy of matter at high stellar densities ($\rho=10^6$ or 10^7 g/cm^3) approaches that due to stellar temperatures (10^6–10^7 °C), it is permissible, for a conservative approximation, to neglect temperatures in computing the rate of combination of protons. The results may then be interpreted as indicating a *minimum* value for the rate of energy production for assumed compositions, or as indicating a maximum hydrogen content for a given rate of energy production.

We will first consider an electron-proton gas, corresponding to a pure hydrogen composition and consider later the modifications necessary to take account of heavier particles. We assume closest cubic packing of the protons. Then, if $2r$ is the distance between nearest protons, the volume per proton is

$$4(\sqrt{2})r^3 \quad \text{and} \quad 2r=(\sqrt{2}M/\rho_H)^{\frac{1}{3}}, \qquad (1)$$

where M is the mass of the proton (1.66×10^{-24} g) and ρ_H is the density of protons in g/cm^3.

2. THE POTENTIAL FUNCTION

In deducing the form of the potential function governing the relative motion of two protons, the influence of neighboring protons will be considered first. The electrical potential at the equilibrium points may be taken as zero. Each proton, in its equilibrium position, is more or

* The preliminary results of this work were referred to by G. Gamow, Phys. Rev. **55**, 723 (1939).

[1] H. Bethe and C. Critchfield, Phys. Rev. **54**, 248 (1938).

TABLE I. *Numerical results for different densities.*

ρ_H	10^4	10^5	10^6	10^7	10^8
$r\times10^{10}$	3.1	1.4	0.66	0.31	0.14
$x_1\times10^3$	1	2.1	4.7	10	21
x_2	0.8	0.75	0.7	0.64	0.57
$E_0(ev)$	22	70	230	750	2500
$\int f(x)dx$	1.29	1.23	1.14	1.02	0.81
G	6×10^{-20}	1.2×10^{-13}	2×10^{-9}	1.9×10^{-6}	3.3×10^{-4}
J	4.4	3.7	3.1	2.7	2.2
Y	0.013	0.018	0.026	0.034	0.052
ϵ/C_H	6×10^{-9}	2.1×10^{-2}	5.8×10^{2}	1.1×10^{6}	4.7×10^{8}

less effectively surrounded by a shell of other protons. According to theory, the potential would be the same at all points within a uniform shell. Actually, the 12 neighboring protons do not form a complete shell, but to a first approximation, there will be no force acting to resist an initial displacement from an equilibrium position.

Imagine two protons to approach each other, one at a time, each moving a distance d from its equilibrium position. In accordance with the foregoing, there will be no change of potential due to the motion of the first. The motion of the second involves a change in potential since its shell is distorted because of the initial displacement of the first proton. The work done in moving the second proton in the field of the first would be $e^2/(2r-2d)-e^2/(2r-d)$ in the absence of the shells. To allow for the effect of the shells, we subtract from the above expression the work which would have been done in moving the second proton in the field of the first proton, had the first not been displaced. This gives the potential due to the protons, as a function of d:

$$V_p(d)=e^2\left(\frac{1}{2r-2d}-\frac{1}{2r-d}-\frac{1}{2r-d}+\frac{1}{2r}\right). \quad (2)$$

Substituting $(r-d)/r=x$ simplifies the expression to

$$V_p(x)=(e^2/2r)[(1-x)^2/x(1+x)]. \quad (3)$$

The kinetic energies of the electrons are very much greater than their interaction with the protons. The electron distribution will therefore not be influenced greatly by the displacement of the protons and the effect of the electrons on the individual protons will be nearly the same as that of a uniformly "smeared out" charge distribution, of density $e/4\sqrt{2}r^3$. For a displacement of a proton by a distance d, there will be a force of

$$\frac{4\pi d^3}{3}\cdot\frac{e^2}{4\sqrt{2}r^3}\cdot\frac{1}{d^2}$$

dynes acting to restore it to its equilibrium position.

The potential due to electrons, for a displacement of *two* protons by a distance d, is, integrating the force expression:

$$V_e(d)=(\pi\sqrt{2}/3)(e^2d^2/2r^3). \quad (4)$$

Since $(d/r)^2=(1-x)^2$, this expression may be combined with (2) to give the total potential

$$V_{(x)}=\frac{e^2}{2r}(1-x)^2\left[\frac{1}{x(1+x)}+\frac{\pi\sqrt{2}}{3}\right]. \quad (5)$$

The approximation involved in assuming that the proton shells and the electrons are not influenced by the motion of an individual proton is on the conservative side, that is, the potential (5) is slightly too large. This will give values of penetration and energy evolution which are slightly low, which is consistent with considering them as lower limits.

3. The Penetration Factor

To obtain the probability of two protons coalescing, we may determine their zero-point energy as an oscillator and use this value of energy in combination with the potential function (5) to calculate the penetration factor. Then, from the probability of the two protons being together, the probability of deuteron formation may be found from Fermi's theory of β-emission, as modified by Gamow and Teller,[2] and Bethe and Critchfield.[1]

The quantum-mechanical formula for the probability of penetration of a potential barrier,[3] when adapted to the notation of Eq. (5), is:

$$G=4\left(\frac{(E-V)_0}{(E-V)_i}\right)^{\frac{1}{2}}\exp\left[-2(2rM)^{\frac{1}{2}}\frac{e}{\hbar}\int_{x_1}^{x_2}\left\{(1-x)^2\left(\frac{1}{x^2+x}+\frac{\sqrt{2}\pi}{3}\right)-E'\right\}^{\frac{1}{2}}dx\right], \quad (6)$$

[2] G. Gamow and E. Teller, Phys. Rev. **49**, 895 (1936).
[3] G. Gamow, *Atomic Nuclei and Nuclear Transformations* (Oxford, 1937), p. 94.

where E', the total energy, is in units of $e^2/2r$. In this expression, $(E-V)_0$ is the kinetic energy outside the barrier; $(E-V)_i$ is the kinetic energy inside the barrier; the lower limit of integration x_1 is the ratio of the width of the potential well to the equilibrium separation; the upper limit x_2 is that value of x for which $E=V_{(x)}$, corresponding to the classical limit of approach.

The proper frequency of vibration in a symmetrical potential of the form (5) is not directly calculable. However, the form of the potential function (5) is sufficiently well approximated, from $x=0.6$ to $x=1$ by the parabola

$$V=2.3e^2(1-x)^2/2r, \tag{7}$$

for which the proper frequency is

$$\nu_0=(2.3e^2/8\pi^2Mr^3)^{\frac{1}{2}}. \tag{8}$$

The zero-point energies $E_0=\frac{1}{2}h\nu_0$ for the values of r corresponding to the different densities, are given on Table I, as are the values of x_2, where $V=E_0$.

The factor $[(E-V)_0/(E-V)_i]^{\frac{1}{2}}$ in (6) may be evaluated from the value of E_0 just obtained. The term $(E-V)_i$ is approximately the depth of the potential hole, 10 Mev, or 16×10^{-6} erg, while $(E-V)_0$ is given by the value of E_0 in Table I. The value of the integral was determined by mechanical integration of the curve plotted from computations. Its values, as well as the values of G finally obtained for various densities are listed in Table I. Each of the variable terms in (6) except the integral have been determined as function of r, and therefore, of ρ_H (by Eq. (1)). To obtain an approximate analytical formula for G, the integral may be replaced by an expression of the type $a(\rho_H)^b$, with a and b determined from the values in Table I for $\rho_H=10^4$ and 10^8. After making the indicated computations, the penetration probability in terms of ρ_H is obtained:

$$G=6.2\times10^{-4}\rho_H^{\frac{1}{2}}\exp\left[-280(\rho_H)^{-0.219}\right]. \tag{9}$$

4. Energy Evolution

The probability of reaction of the two protons (the probability of β-emission per second) is given by Bethe and Critchfield as

$$\beta=1.2\times10^{-5}M^2\ \text{sec.}^{-1}, \tag{10}$$

where

$$M=|\psi_p\psi_d d\tau| \tag{11}$$

with ψ_p and ψ_d the normalized wave functions of protons and deuteron, respectively, and $d\tau$ the space element. Their detailed calculations of ψ_p and ψ_d are not directly applicable here. As a conservative approximation, we assume the simplified radial wave functions:

$$\psi_d=r_0^{-\frac{1}{2}},\quad \psi_p=[r(1-x_2)]^{-\frac{1}{2}}G^{\frac{1}{2}} \tag{12}$$

inside the deuteron potential hole, with $\psi_d=0$ outside, and integrate over the radius of the potential hole r_0, which is taken as 3×10^{-13} cm. This gives:

$$\beta=1.2\times10^{-5}Gr_0/r(1-x_2) \tag{13}$$

for protons in an S state, which is the state implied by the form of Eqs. (12).

The maximum angular momenta computed from the zero-point energies and the proton separations, which are given in units of $\hbar$ as J in Table I, indicate that not all pairs of neighboring protons can be considered as in S states. Since the probability of penetration is much less for other states, we neglect it entirely. We take the probability of a given pair of protons having relative angular momentum less than $\hbar/2$ as the probability of their being in an S state. This is approximately

$$Y=(1/2J)^2=\hbar^2/16r^2ME_0, \tag{14}$$

the values of which are listed in Table I.

Since there are twelve nearest neighbors for each particle in the close packed cubic lattice, the probability of reaction will be increased, but only by a factor of six, since in summing, each pair of neighboring protons must be counted only once. Because of the influence of the spin on penetration probabilities, only those protons

TABLE II. *Minimum energy evolution for different densities and hydrogen concentration.*

Density (g/cm³)		10^6	5×10^6	10^7	5×10^7
ϵ_0 for $C_H=1$, (erg/g sec.)	K	1.4×10^{-9}	0.26	82	1.3×10^4
	W	6×10^2	1×10^5	1×10^6	3×10^7
ϵ_0 for $C_H=0.3$, (erg/g sec.)	K	1.2×10^{-19}	1.3×10^{-7}	4.9×10^{-4}	4.3×10^2
	W	8	2×10^3	2×10^4	6×10^5

with opposite spins may be expected to react, and this reduces the probability by a factor of $\frac{1}{4}$. In view of the approximations above, these opposing factors might be neglected, but we retain them for completeness.

Each proton-proton combination initiates a chain of reactions leading probably to the formation of a $_2He^4$ nucleus with the evolution of 4.3×10^{-5} erg.[1] However, Alvarez and Cornog[4] have found that $_2He^3$ is a stable isotope of helium. If $_2He^3$ is the end product, the energy evolution will be somewhat smaller. From the value given with Eq. (13), the energy production *per gram* (6×10^{23} protons for a composition of 100 percent hydrogen) is

$$\epsilon_H=4.6\times10^{14}GYr_0/r(1-x_2). \qquad (15)$$

There seems to be no reason to expect that the reaction mechanism will be much changed by the inclusion of other nuclei besides protons. The probability of a given proton reacting might be slightly reduced by heavier nuclei among its nearest neighbors, but this effect would reduce the energy evolution only in proportion to the number of protons displaced, say by $\frac{1}{3}$ for reduction of the hydrogen content from 100 percent to 40 percent. On the other hand, because of the stronger repulsion by the heavier nuclei, zero-point energies of the protons would increase. From this point of view, (15) may be expected to give a *minimum* value for ϵ_H independent of total density as long as the hydrogen content is at all appreciable, and is to be interpreted as giving the rate of energy evolution *per gram of hydrogen*. The energy production rate per gram of total stellar material would be

$$\epsilon=4.6\times10^{14}GYr_0C_H/r(1-x_2), \qquad (16)$$

where C_H is the percentage, by weight, of hydrogen, r is given by Eq. (1) with $\rho_H=\rho C_H$ and G and x_2 correspond to this value of r. The rates of energy production computed on this formula are given in Table I for different values of ρ_H. Substituting the expressions (1) (9) and (14) for the various factors, in terms of $\rho_H=\rho C_H$, we obtain the approximate analytical formula:

$$\epsilon=4.3\times10^{5}\rho^{\frac{1}{2}}C_H^{5/3}\exp\left[-280(\rho C_H)^{-0.22}\right]. \qquad (17)$$

[4] Alvarez and Cornog. Phys. Rev. **56**, 379 (1939).

5. Kothari's Formula

While this investigation was in progress, a note appeared by Kothari,[5] on the lower limit of energy evolution in stellar matter. He applied the Bethe-Critchfield formula to degenerate protons, indicating that proper changes had been made to take account of the new velocity-distribution law. His final formula was

$$\epsilon_0=2.4\times10^{3}\rho C_H^2e^{-z}, \qquad (18)$$

where C_H is the hydrogen concentration, and

$$z=4.2\times10^{3}(\rho C_H)^{-\frac{1}{3}}.$$

The rates of energy production calculated by the two methods are at extreme variance. Kothari's results are given in Table II, with our (approximate) corresponding values.

It may be instructive to consider briefly in a semi-quantitative fashion the modification of the Bethe-Critchfield treatment necessary to take account of the Fermi-Dirac distribution and to compare the resulting energy formula with (17) and (18). Bethe and Critchfield[1] developed their formula from the general equation for the number of combinations per cm^3 per sec., p', (their Eq. (30)):

$$p'=(1/4N^2)(V)dVv\sigma, \qquad (19)$$

where N is the number of protons per cm^3, (V) the distribution function for the relative velocity V, v the scalar velocity, and σ the cross section for combination of two protons of relative velocity V. From this point on, their development included the evaluation of the integral for a Boltzmann distribution function. Presumably Kothari diverges at this point to obtain an evaluation of the integral for the relative velocities in a Fermi-Dirac gas, using the same expression for σ obtained by Bethe and Critchfield on the basis of Coulomb potential barriers.

The distribution function in (19) may be replaced by a delta-function, which is unity for the average effective velocity and zero for all other values of V. Without deriving the value of the average effective velocity here, it will obviously be less than the maximum relative velocity, which is, from the Fermi-Dirac formula:

$$2V_0=2(h/m)(3N/8\pi)^{\frac{1}{3}}, \qquad (20)$$

[5] D. S. Kothari, Nature **142**, 916 (1938).

TABLE III. *Observed and computed rates of energy production in white dwarf stars.*

STAR	CENTRAL DENSITY G/CM³	ϵ OBS.	ϵ (CALC.; $C_H=0.3$)
Sirius B	10^6	6×10^{-3}	10
O_2 Eridani	5×10^5	3×10^{-3}	10^{-1}
Van Maanen's	10^7	1×10^{-4}	10^4
A C 70° 8247	5×10^7	2×10^{-2}	10^5

so that

$$p' < \tfrac{1}{4}N^2(2V_0)\sigma, \tag{21}$$

where σ is the cross section for the relative velocity $2V_0$. The energy evolution per gram per second is given by

$$\epsilon = 4.3\times10^{-5}p'/\rho. \tag{22}$$

From Bethe and Critchfield, for $V=2V_0$

$$\sigma = (3\times10^{-31}/4V_0^2)\exp\left[-1.45\times10^9/2V_0\right]. \tag{23}$$

Since the velocity appears in the exponent and the faster protons are the more numerous in the Fermi-Dirac distribution, the average effective velocity will not be greatly different from $2V_0$ and the right side of (21) may be expected to give p' to the correct order of magnitude.

Substituting (20) and (23) in (22) with $N=\rho C_H\times6\times10^{23}$, we have

$$\epsilon = 3.4\times10^6\rho^{\frac{1}{2}}C_H^{5/3}\exp\left[-4.26\times10^3(\rho C_H)^{-\frac{1}{3}}\right] \tag{24}$$

as an expression setting an upper limit for the energy evolution on the assumption of Fermi-Dirac energy values and cross sections based on a Coulomb potential. It may be noted that the exponent in (24) is precisely the same as in Kothari's equation, (18). The coefficient is greater than Kothari's by a factor of $1400/(\rho C_H)^{\frac{1}{2}}$ which for $\rho C_H=10^7$ is only about 7.

Comparing (18) and (24) with (17) it is seen that the coefficient in (17) differs from Kothari's by a factor of 178 $(\rho C_H)^{-\frac{1}{2}}$, or 1.78 at $\rho C_H=10^6$. In the exponent, C_H enters to the $\frac{1}{3}$ power on the basis of the Fermi-energy and the Coulomb potential, but only to the 0.22 power in our development. Not only does the exponent vary more slowly with ρC_H in (17) than in (18) or (24), but it is only about one-third as large at $\rho C_H=10^6$.

The Bethe-Critchfield formula was derived on the basis of a free path long in comparison with the effective radius of the protons—i.e., by using a Coulomb potential. But (5) shows that the potential differs radically from the Coulomb form when the nearness of other protons at high densities is taken into account. If we equate the Coulomb potential energy $e^2/r=3.1\times10^{-11}(\rho C_H)^{\frac{1}{2}}$ to the maximum value of the zero-point energy for the Fermi-Dirac proton gas,

$$E_0 = 2\times10^{-14}(\rho C_H)^{\frac{1}{2}},$$

we obtain

$$\rho C_H > 3\times10^9 \tag{25}$$

as the condition for proton energies to be great enough to allow two protons to approach each other within their actual separation distances. Thus at $\rho C_H=10^6$, for instance, if we assume a Coulomb potential between two protons and neglect the effect of electrons, we have the paradoxical situation that the Fermi energy will permit an approach to only 5×10^{-10} cm, while the average separation distance is 1.2×10^{-10} cm! This paradox disappears at high temperatures and low densities; the Bethe-Critchfield assumptions are quite valid for the densities and temperatures which they considered ($\rho\sim10^2$; $T\sim10^7$). However, it indicates strikingly the burden placed on the Fermi-Dirac energies by using the unmodified potential at $\rho=10^6$ to 10^7.

It seems also that our procedure of quantizing the motion of the protons in the modified potential (5) is preferable to the Fermi-Dirac formula for point particles which takes no account of the effective finite size of the particles due to the thickness of the electrostatic potential barrier. The effect of the barrier thickness is to decrease the effective separation, or free path, and therefore also the wave-length, thus increasing the energy and greatly increasing the penetration.

However, the Fermi energy increases with $(\rho C_H)^{\frac{1}{2}}$, while the zero-point energy of an oscillator between potential barriers of the form (5) increases roughly as $(\rho C_H)^{0.5}$ over the range of densities considered. Thus it appears that at higher densities ($\rho C_H\sim10^9$), the greater increase of the Fermi energy will cause the values of ϵ computed by (18) to equal those computed by (17). As indicated by the inequality (25), it is just in this range that the zero-point energies become so great that the form of the lower part of the potential barrier is of no great consequence.

At even greater densities, it is probable that Kothari's formula is to be preferred, but these are beyond the range which is of special physical interest.

6. Astrophysical Consequences

The significance of the results is apparent in the determination of the composition of white dwarf stars. Relevant data on a few of these stars are given in Table III, based on values given by Chandrasekhar.[6] The rates of energy production given in the table were computed on the basis of the central densities. It is seen that the computed values are much greater than the values obtained from observational data. The conclusion is that only a very inconsiderable part of most of these stars can be composed of hydrogen.

This conclusion lends support to the theories of stellar evolution which place the white dwarfs at the end of the evolutionary scale. It also necessitates a critical review of either the observational data on these stars or their theoretical interpretation. Without considering in detail the equations of structure of degenerate stars, it will suffice to recall that certain relations (depending somewhat on the model chosen) subsist among the variables: mass, luminosity, radius, and effective molecular weight. For degenerate stellar material, the effective molecular weight, μ_e is the mass, in atomic units, per *electron*. For hydrogen μ_e is unity; for $_2He^4$ and heavier light elements, 2.0; for heavy elements, 2.0 to about 2.5. The heavier elements presumably occur only in small quantities, so that it seems to have been assumed that there are essentially only two different values for μ_e, namely, 1.0 for hydrogen, 2.0 for all other elements. Thus, when the other variables are determined from observational data, the value of μ_e obtained from the equations of structure is used to determine C_H, the relation between μ_e and C_H being approximately

$$C_H = 2.0 - \mu_e. \tag{26}$$

The values of C_H so obtained for Sirius B, (the white dwarf for which the observational data are the most reliable) are rather large; Chandrasekhar (reference 6, p. 433) gives $C_H = 0.52$, and Eddington[7] gives $C_H = 0.66$.

It seems clear that these values are untenable in the light of the preceding calculations; the energy production for such concentrations would be vastly larger than those corresponding to the observational data. On the other hand, it does not seem that the observational data on this star are likely to be greatly enough in error to overcome this difficulty.

One possible escape from the paradox is suggested by the work of Alvarez and Cornog[5] who separated a beam of $_2He^3$ from ordinary helium, using the 60-inch cyclotron. While the abundance ratio of this isotope is apparently very small compared to $_2He^4$, the fact that it is stable under terrestrial conditions suggests the possibility that it may be stable under stellar conditions, also.

The reaction

$$H + H = D + e^+ \tag{27}$$

is quickly followed by the reaction

$$D + H = {}_2He^3 + \gamma. \tag{28}$$

However, the further reaction assumed by Bethe and Critchfield:

$$_2He^3 + {}_2He^4 \pm {}_4Be^7 + \gamma \tag{29}$$

might proceed slowly enough for fairly large amounts of $_2He^3$ to be present.

The value of μ_e for $_2He^3$ is 1.5. Thus even if it is assumed that no hydrogen remains, values of μ_e less than 2.0 (down to 1.5) may still be explained as due to appropriate amounts of light helium. There are serious objections to this possibility for it is difficult to see how the star could arrive at the white dwarf stage without passing through an earlier stage of higher temperatures which would have furthered the reaction (29). It appears, therefore, that the difficulties can be overcome only by the revision of the values of the white dwarf radii, as obtained from the observational data.

In conclusion, I am happy to express my great indebtedness to Dr. G. Gamow, for suggesting the problem, to Dr. E. Teller for valuable suggestions and careful criticism of the manuscript, and to both for stimulating discussions.

[6] S. Chandrasekhar, *An Introduction to the Study of Stellar Structure* (University of Chicago Press), Chap. 11.

[7] A. S. Eddington, Monthly Notice **99**, 595–606 (1939).

Relation Between Actual and Virtual Ionospheric Height

H. G. Booker and S. L. Seaton
Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D. C.
(Received August 24, 1939)

Translation of virtual height of the ionosphere into actual height is achieved by fitting a parabolic maximum of electron density to observations of variation of virtual height with wave frequency. Methods for measuring the thickness of a region and the height of maximum electron density are described and applied. The simplest measure of the actual height of maximum electron density is the virtual height at five-sixths of the penetration frequency. This measure is found to be remarkably reliable for F region at night. During the daytime it is often necessary to correct F_1 observations for presence of E region, and F_2 observations for presence of F_1 region. A simple technique of correction is described and applied. The disadvantages of minimum virtual height for quantitative study of diurnal variation of ionospheric height are emphasized.

I. Introduction

THE object of this paper is to describe a method for translating virtual height of the ionosphere into actual height. The method is quite simple, and yet sufficiently accurate for many practical purposes. It applies primarily to the F region, but difficulty arises when this region is separated into F_1 and F_2 regions. It is assumed that the observational material consists of variation with frequency of virtual height of the ionosphere for radio waves reflected vertically. Virtual height is proportional to time delay. Observations of this type over practically the entire range of wave frequency for which echoes are normally received from the F region are made automatically every 15 minutes at the observatories maintained at Huancayo, Peru, and Watheroo, Western Australia, by the Department of Terrestrial Magnetism of the Carnegie Institution of Washington. The problem of translating virtual height into actual height is one to which thought has already been given.[1-3] One method is to fit to the observations a parabolic maximum of electron density (see Fig. 1). It is this method which will be considered here.

In Section II we treat an idealized problem. It is supposed that: (i) The ionosphere consists of a parabolic maximum of electron density; (ii) propagation of radio waves is unaffected by presence of the earth's magnetic field; (iii) there is no Lorentz polarization correction;[4,5] (iv) observations have been made of variation of virtual height with wave frequency from zero frequency up to the penetration frequency. The problem is to find convenient methods for deducing from the observations the constants which define the parabolic maximum of electron density. The extent to which the results of Section II are applicable in practice is discussed in Section III, and the actual distribution of electron density with height is deduced for two sample ionospheric records. In Section IV the results of Sections II and III are considered in connection with the

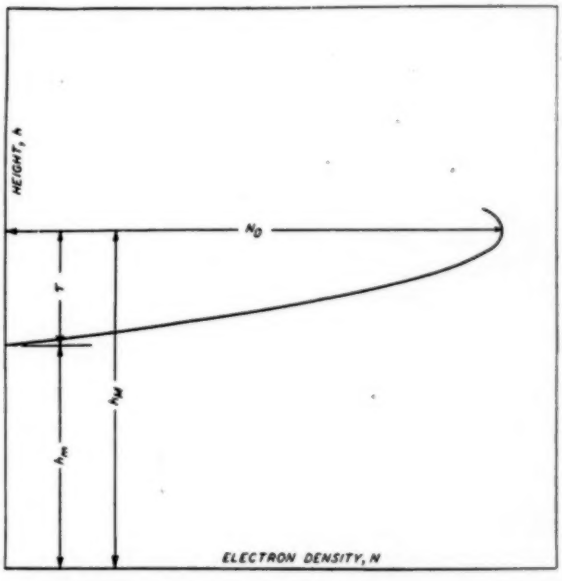

FIG. 1. Parabolic maximum of electron density.

[1] E. V. Appleton, Proc. Phys. Soc. **41**, 43–59 (1928).
[2] E. V. Appleton, Proc. Roy. Soc. **A162**, 451–479 (1937).
[3] F. H. Murray and J. B. Hoag, Phys. Rev. **51**, 333–341 (1937).
[4] C. G. Darwin, Proc. Roy. Soc. **A146**, 17–46 (1934).
[5] H. G. Booker and L. V. Berkner, Terr. Mag. **43**, 427–450 (1938).

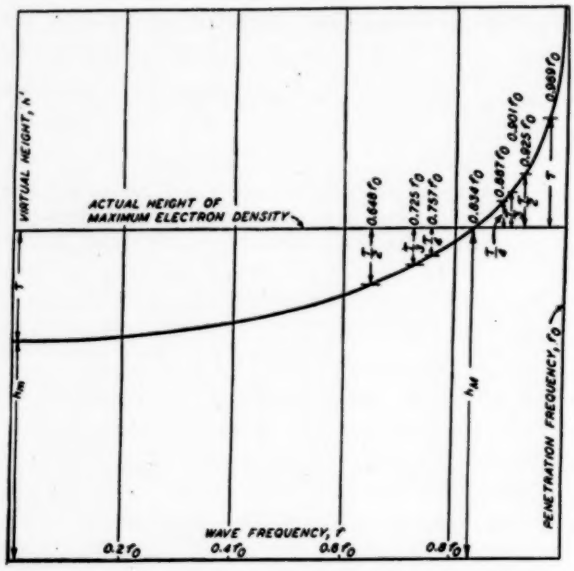

FIG. 2. Relation between virtual height and wave frequency for parabolic maximum of electron-density.

problem of routine scaling of a mass of ionospheric records.

II. PARABOLIC MAXIMUM OF ELECTRON DENSITY

An ionosphere consisting of a parabolic maximum of electron density is depicted in Fig. 1. The distribution of electron-density N with height h is defined by three quantities: (i) The maximum electron density, N_0; (ii) the height h_M at which the electron density is maximum; (iii) the height h_m at which the electron density begins to increase from zero.

The relation between N and h is

$$\frac{N_0-N}{N_0}=\left(\frac{h_M-h}{h_M-h_m}\right)^2. \tag{1}$$

The quantity

$$\tau=h_M-h_m \tag{2}$$

will be called the thickness of the parabolic maximum of electron density.

An ionosphere of the type shown in Fig. 1 has a penetration-frequency f_0. Only radio waves of frequency less than f_0 are reflected at vertical incidence. It is well known [compare reference 2] that, on the assumptions mentioned in Section I, the variation of virtual height h' with wave frequency f for the parabolic maximum (1) of electron density is given by

$$\frac{h'-h_m}{h_M-h_m}=\frac{f}{2f_0}\log_e\frac{f_0+f}{f_0-f}. \tag{3}$$

This variation is depicted in Fig. 2.

The problem is: Given the variation of h' with f shown in Fig. 2, to determine the corresponding variation of N with h shown in Fig. 1. This is equivalent to determining the three quantities N_0, h_M, and h_m, or alternatively the

TABLE I. *Tabulation of* $\varphi(x)=\frac{x}{2}\log_e\left|\frac{1+x}{1-x}\right|-1$.

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	−1.000	−1.000	−1.000	−0.999	−0.998	−0.997	−0.996	−0.995	−0.994	−0.992
0.1	−0.990	−0.988	−0.986	−0.983	−0.980	−0.977	−0.974	−0.971	−0.967	−0.963
0.2	−0.959	−0.955	−0.951	−0.946	−0.941	−0.936	−0.931	−0.925	−0.919	−0.913
0.3	−0.907	−0.901	−0.894	−0.887	−0.880	−0.872	−0.864	−0.856	−0.848	−0.839
0.4	−0.830	−0.821	−0.812	−0.802	−0.792	−0.782	−0.771	−0.760	−0.749	−0.737
0.5	−0.725	−0.713	−0.700	−0.687	−0.674	−0.660	−0.646	−0.631	−0.616	−0.600
0.6	−0.584	−0.568	−0.551	−0.533	−0.515	−0.496	−0.477	−0.457	−0.436	−0.415
0.7	−0.393	−0.370	−0.347	−0.322	−0.297	−0.270	−0.243	−0.214	−0.185	−0.154
0.8	−0.121	−0.087	−0.051	−0.014	+0.026	+0.068	+0.112	+0.160	+0.211	+0.265
0.9	+0.325	+0.390	+0.462	+0.542	+0.634	+0.740	+0.868	+1.029	+1.252	+1.620
1.0	$+\infty$	+1.678	+1.354	+1.170	+1.045	+0.950	+0.874	+0.812	+0.759	+0.714
1.1	+0.674	+0.639	+0.608	+0.580	+0.554	+0.531	+0.510	+0.490	+0.471	+0.455
1.2	+0.439	+0.424	+0.410	+0.397	+0.385	+0.373	+0.362	+0.352	+0.342	+0.333
1.3	+0.324	+0.316	+0.307	+0.300	+0.292	+0.285	+0.279	+0.272	+0.266	+0.260
1.4	+0.254	+0.249	+0.243	+0.238	+0.233	+0.229	+0.224	+0.220	+0.215	+0.211
1.5	+0.207	+0.203	+0.199	+0.196	+0.192	+0.189	+0.186	+0.182	+0.179	+0.176
1.6	+0.173	+0.170	+0.167	+0.165	+0.162	+0.159	+0.157	+0.155	+0.152	+0.150
1.7	+0.148	+0.145	+0.143	+0.141	+0.139	+0.137	+0.135	+0.133	+0.131	+0.129
1.8	+0.127	+0.126	+0.124	+0.122	+0.120	+0.119	+0.118	+0.116	+0.114	+0.113
1.9	+0.112	+0.110	+0.109	+0.107	+0.106	+0.105	+0.104	+0.102	+0.101	+0.100

three quantities N_0, h_M, and τ. N is derived directly from the observed penetration frequency f_0. If N_0 is measured per cc and f_0 in mc/sec., then

$$N=kf_0^2, \tag{4}$$

where

$$k=1.24\times10^4. \tag{5}$$

The problem is therefore reduced to the following: Given the variation of h' with f shown in Fig. 2, to deduce h_M and h_m, or h_M and τ.

Define a function

$$\varphi(x)=\frac{x}{2}\log_e\left|\frac{1+x}{1-x}\right|-1. \tag{6}$$

$\varphi(x)$ is plotted in Fig. 3 and tabulated in Table I. Using (2) and (6), one may write (3)

$$h'=h_M+\tau\cdot\varphi(f/f_0). \tag{7}$$

Given the variation of h' with f, we may deduce h_M and τ as follows: (i) Measure the values of h' corresponding to any two values of f/f_0 lying between 0 and 1; (ii) from Table I find the values of φ corresponding to these two values of f/f_0; (iii) substitute the two pairs of corresponding values of h' and φ into (7), and solve for h_M and τ.

In practice it is convenient to choose the two values of f/f_0, not at random, but so that the corresponding values of φ are numerically simple. Values of f/f_0 corresponding to a series of useful values of φ are tabulated in Table II and indicated in Fig. 2. The values in Table II are deduced from Table I.

Consider the wave frequency

$$f=0.834f_0. \tag{8}$$

From Table II the corresponding value of φ is zero. Hence, at the wave frequency (8), (7) becomes

$$h'=h_M. \tag{9}$$

Eqs. (8) and (9) state that, at the wave frequency equal to 0.834 times the penetration frequency, the virtual height is the actual height of maximum electron density. Thus, if h'_{834} is the virtual height at the wave frequency equal to 0.834 times the penetration frequency, then

$$h_M=h'_{834} \tag{10}$$

(see Fig. 2). This remarkable result is valid no matter what may be the values of the thickness τ and the penetration frequency f_0. The peculiar property of the wave frequency (8) is demonstrated in Fig. 4, which shows the effect upon the (h',f) curve of varying τ keeping h_M constant. We may note that 0.834 (or more accurately 0.8336) is practically five-sixths.

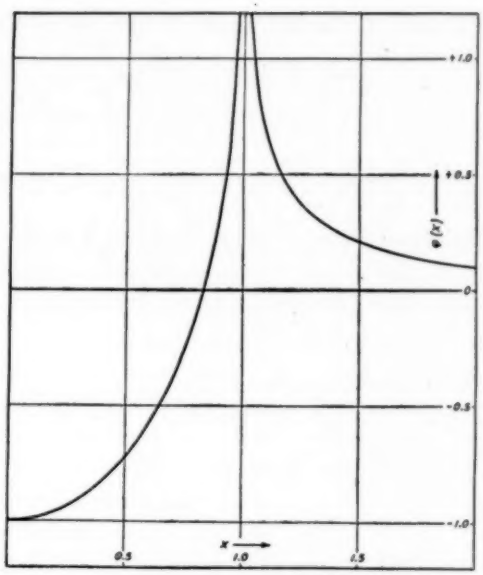

FIG. 3. Graph of $\varphi(x)=x/2\left(\log_e\left|\frac{1+x}{1-x}\right|-1\right)$.

Measurement of h' at any two of the wave frequencies mentioned in Table II gives immediately the values of both h_M and τ. For example, let h'_{648} and h'_{925} be the virtual heights at the wave frequencies equal to 0.648 and 0.925 times the penetration frequency. From Table II the values of φ corresponding to these wave frequencies are $-\frac{1}{2}$ and $+\frac{1}{2}$, respectively. Hence (7) gives

$$\begin{cases} h'_{648}=h_M-(1/2)\tau, & (11)\\ h'_{925}=h_M+(1/2)\tau. & (12)\end{cases}$$

TABLE II. *Useful wave frequencies.*

$\varphi=$	-1	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	1
$f/f_0=$	0	0.648	0.725	0.757	0.834	0.887	0.901	0.925	0.969

Solving (11) and (12) for h_M and τ we obtain

$$\begin{cases} h_M=(1/2)(h'_{925}+h'_{648}), & (13) \\ \tau=h'_{925}-h'_{648}. & (14) \end{cases}$$

The validity of (13) and (14) is also obvious from Fig. 2.

Formulas for h_M and τ involving values of h' at more than two wave frequencies are easily deduced. For example, it is obvious from Fig. 2 that

$$\begin{cases} h_M=(1/5)\{(h'_{925}+h'_{648}) \\ \qquad +(h'_{887}+h'_{757})+h'_{834}\}, & (15) \\ \tau=(4/5)(h'_{925}-h'_{648}) \\ \qquad +(2/5)(h'_{887}-h'_{757}). & (16) \end{cases}$$

III. Distribution of Electron Density with Height

Figure 5 shows a typical example of an ionospheric record obtained at night by the automatic multifrequency technique.[6] Between sunset and sunrise the ionosphere consists normally of a single main ionized region—the F region. We shall show that the results of Section II may be applied directly to records such as this to quite a high degree of accuracy. To do this we must examine the validity of the assumptions upon which Section II is based.

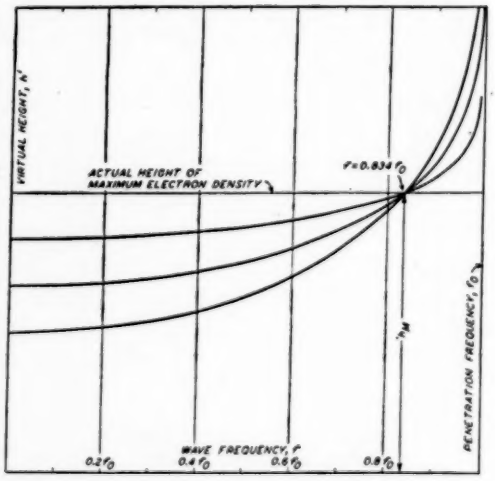

Fig. 4. Effect upon (h',f) curve of varying τ keeping h_M constant.

[6] L. V. Berkner, H. W. Wells and S. L. Seaton, Trans. Edinburgh Meeting 1936; Internat. Union Geod. Geophys., Ass. Terr. Mag. Electr., Bull. No. 10, pp. 340–357 (1937).

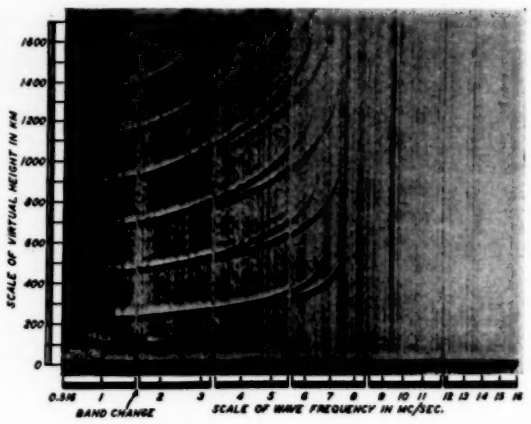

Fig. 5. Observed variation of virtual height with wave frequency, Watheroo magnetic observatory, March 14, 1939, 20h 15m to 20h 30m local time (120° East Meridian).

It is assumed that the ionosphere consists of a parabolic maximum of electron density. According to Section II h_M and τ may be deduced from the observed values of h' at any two wave-frequencies. Table III shows the results obtained for the record reproduced in Fig. 5 with pairs of wave frequencies mentioned in Table II. Heights were scaled by use of the second multiple echo ($2F$). The consistency of results obtained with the ordinary wave is quite remarkable. As might be expected, results were not quite so consistent for the extraordinary wave. There is little doubt that, for the record reproduced in Fig. 5, the distribution of electron density with height is substantially parabolic as shown in Fig. 1. This distribution is maintained from the maximum down to an electron density, say, 20 percent of the maximum. However, as might be expected, it is not maintained down to zero electron density. This is demonstrated by the fact that the minimum virtual height of the F region for the record reproduced in Fig. 5 is only 225 km, whereas h_m is 247 km.

It is assumed in Section II that propagation of radio waves in the ionosphere is unaffected by the

Table III. *Consistency of deduced heights.*

Ratio of Frequencies Used to Penetration Frequency	Ordinary Wave			Extraordinary Wave		
	h_M	h_m	τ	h_M	h_m	τ
0.648, 0.925	359	247	112	350	240	110
0.725, 0.901	358	247	111	350	230	120
0.757, 0.887	358	248	110	355	235	120
0.834, 0.969	358	246	112	355	220	135

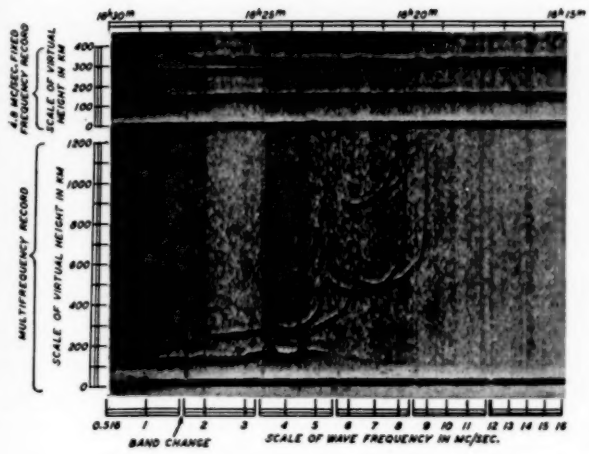
4.8 MC/SEC. FIXED FREQUENCY RECORD
SCALE OF VIRTUAL HEIGHT IN KM
MULTIFREQUENCY RECORD
SCALE OF VIRTUAL HEIGHT IN KM
BAND CHANGE
SCALE OF WAVE FREQUENCY IN MC/SEC.

FIG. 6. Observed variation of virtual height with wave frequency, Watheroo magnetic observatory, December 18, 1939, $16^h\ 15^m$ to $16^h\ 30^m$ local time (120° East Meridian).

earth's magnetic field. At Huancayo, close to the geomagnetic equator, this assumption may be regarded as fully satisfied for the ordinary wave. But at Watheroo, where the magnetic dip is about $-64°$, the earth's magnetic field affects propagation of both magneto-ionic components. However, the work of Goubau[7] shows that the correction required on this account is at least twice as great for the extraordinary wave as for the ordinary. It follows that, first, it is better to apply the methods of Section II to the ordinary wave rather than the extraordinary wave and, second, the error in so doing should be less than the difference in results obtained from the two magneto-ionic components. The difference between the virtual heights for the two components at the wave frequencies equal to 0.834 times their respective penetration frequencies is usually between five and ten km when the ionosphere is substantially a parabolic maximum of electron density. We conclude that effect of the earth's magnetic field upon results obtained from the ordinary wave is small. Nevertheless it is systematic.

Similar remarks apply to the assumption in Section II that there is no Lorentz polarization correction. The only important effect of including this correction would be to replace (5) by

$$k = 1.86 \times 10^4. \qquad (17)$$

[7] G. Goubau, Hoch: tech. u. Elek: akus. **44**, 17–23, 138–139 (1934).

We may sum up by saying that, for records of the type reproduced in Fig. 5, the technique of Section II applied to the ordinary wave gives results sufficiently accurate for most practical purposes.

Between sunrise and sunset, records are not usually as simple as that reproduced in Fig. 5. Fig. 6 shows a record (the trace at the top of this record shows variation of virtual height with time for a fixed wave frequency of 4.8 mc/sec.) more typical of conditions existing during daylight. Three main regions are apparent—E, F_1, and F_2. In these circumstances it is clearly impossible to represent the ionosphere by a single parabolic maximum of electron density. It is possible, however, to represent it by a triplet of parabolas as shown in Fig. 7. Fig. 7 was deduced from the trace corresponding to the ordinary wave on the record reproduced in Fig. 6 in the following way:

(i) The technique of Section II was first applied to E region.

(ii) The E region was assumed to consist of a complete parabolic maximum of electron density and observations of virtual height for F_1 region were corrected for presence of E region. This was done by subtracting from the virtual height of F_1 region at wave frequency f an amount $2\tau \cdot \varphi(f/f_0)$, where τ and f_0 are, respectively, the thickness and penetration frequency deduced for E region, and the function φ is given by Table I.

(iii) The technique of Section II was applied to F_1 region corrected as above.

(iv) The F_1 region was assumed to consist of half of a parabolic maximum of electron density

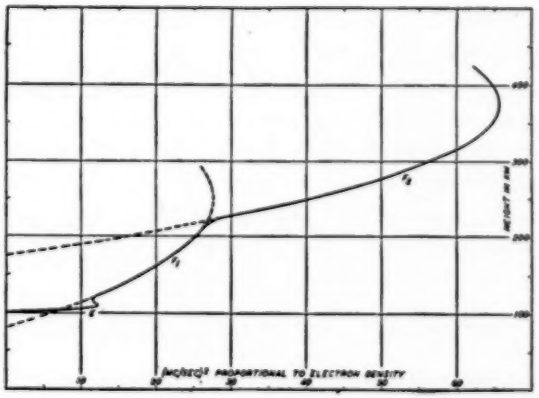

FIG. 7. Distribution of electron density with height corresponding to record reproduced in Fig. 6.

FIG. 8

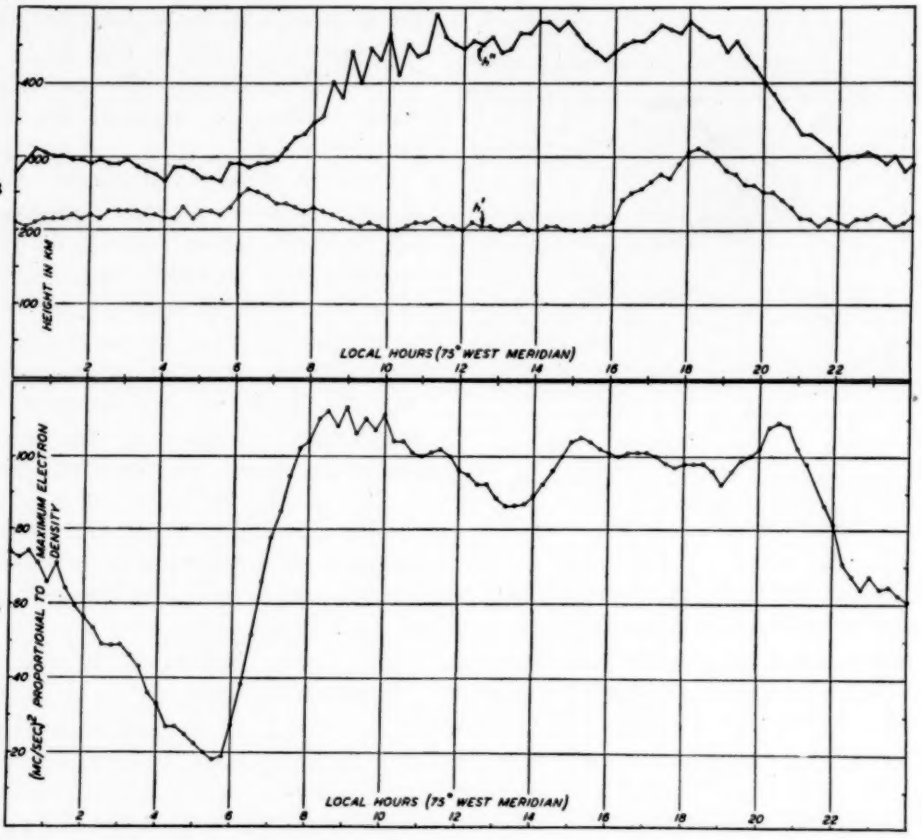

FIG. 9.

FIG. 8. Diurnal variation of h' and h'' for F region, Huancayo magnetic observatory, May 31, 1939.
FIG. 9. Diurnal variation of square of penetration frequency of F region for ordinary wave, Huancayo magnetic observatory, May 31, 1939.

and observations of virtual height for F_2 region were corrected for presence of F_1 region. This was done by subtracting from the virtual height of F_2 region at wave frequency f an amount $\tau \cdot \varphi(f/f_0)$, where τ and f_0 are, respectively, the thickness and penetration frequency deduced for F_1 region. Correction of F_2 observations for presence of E region was found to be negligible.

(v) The technique of Section II was applied to F_2 region corrected as above.

It is believed that Fig. 7 is quite a good representation of the distribution of electron density with height corresponding to the record reproduced in Fig. 6. Note the low actual heights associated with F_1 region.

IV. Routine Scaling of Ionospheric Records

In mass analysis of ionospheric records it has been the practice to make scalings for E, F_1, and F_2 (or F) regions as follows. Records made at hourly intervals are scaled for penetration frequency f_0 of the ordinary wave and for minimum virtual height. It would now be desirable to replace scaling of minimum virtual height by scaling of the height h_M of maximum electron density and the thickness τ by the technique of Sections II and III. This would provide practically a complete record of the distribution of electron density with height in the ionosphere up to the level of maximum electron density. Moreover, it would then be possible to study variation of the quantity $(2/3)\tau k f_0^2$ which is representative of the total number of electrons in a region below the level of maximum electron density. Such a program may not be practical because of the labor involved. In the meantime it is desirable that a less ambitious program be drawn up.

Consider the following program. Scale for the F_1 and F_2 (or F) regions: (i) The penetration

FIG. 10.

FIG. 11.

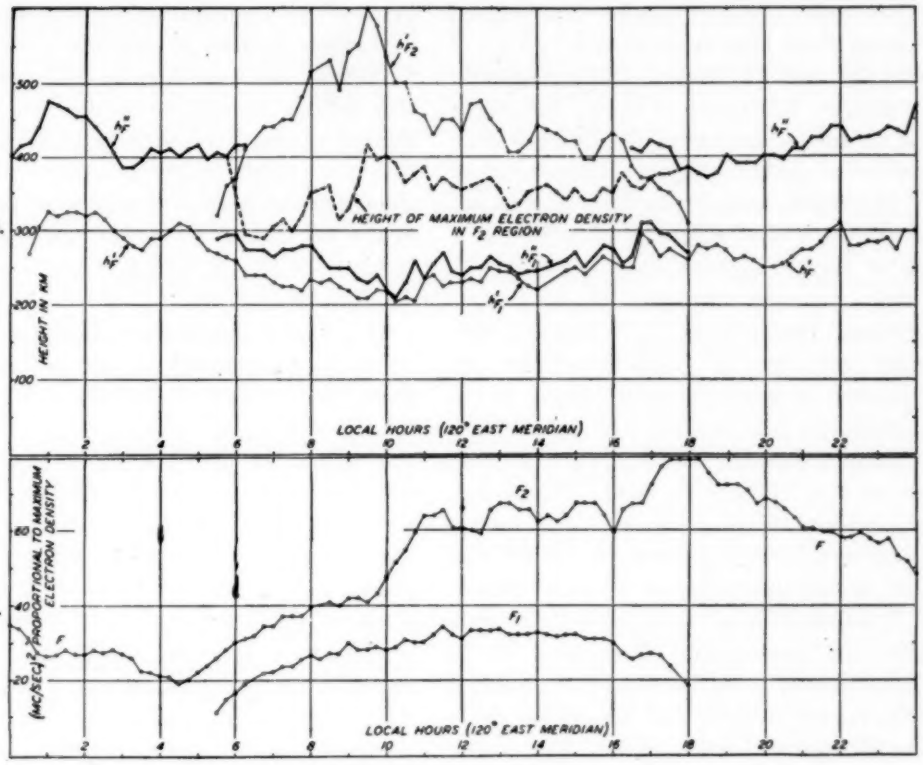

FIG. 10. Diurnal variation of h' and h'' for F_1 and F_2 regions, Watheroo magnetic observatory, December 18, 1938.
FIG. 11. Diurnal variation of squares of penetration frequencies of F_1 and F_2 regions for ordinary wave, Watheroo magnetic observatory, December 18, 1938.

frequency f_0 for the ordinary wave; (ii) the virtual height h'' for the ordinary wave at the wave frequency $0.834f_0$; (iii) the minimum virtual height h'. The result of scaling in this way the records from Huancayo for a single day is shown in Figs. 8 and 9. Separation of F region during daylight into F_1 and F_2 regions was not sufficiently marked to take account of. At night h' does not differ greatly from h_m (Fig. 1), and gives a useful indication of what may be called the bottom of F region. During the day, however, h' is appreciably greater than h_m owing to presence of E region. h'' is quite a satisfactory measure of the height h_M of maximum electron density in F region throughout this particular day. A check on this was obtained by comparing two values for the intensity of the earth's magnetic field. One is the intensity deduced for height h'' from the intensity at ground level on the assumption that the strength of the earth's magnetic field varies inversely as the cube of distance from the center of the earth. The other is the intensity deduced for the actual level of maximum electron density from the penetration frequencies of the region for the ordinary and extraordinary waves.[8] The average of the 96 ratios of the second intensity to the first obtained for the day represented in Figs. 8 and 9 is 0.99 ± 0.01. This is to be regarded as a satisfactory check. It may be mentioned that it is difficult to obtain a similar check at Watheroo for the following reason. Because of the inclination of the earth's magnetic field at Watheroo, the ordinary wave is deviated in the magnetic meridian towards the south pole and the extraordinary wave towards the equator. Consequently, the two magneto-ionic components do not penetrate the ionosphere at exactly the same points. There is a small systematic difference between the maximum electron densities at the two points of penetration, and this is sufficient to ruin the check unless a reliable correction can be applied.

[8] E. V. Appleton, Nature **133**, 793 (1934).

Figures 10 and 11 show a more complicated state of affairs than that represented in Figs. 8 and 9. Figs. 10 and 11 refer to a magnetically disturbed day at Watheroo. During the night h_F' gives a useful indication of the bottom of F region, and h_F'' is quite a satisfactory measure of the height of maximum electron density. During the day, however, the situation is very different. At sunrise F region separates into F_1 and F_2 regions. h_{F2}' increases rapidly to enormous virtual heights. Between $06^h\,15^m$ and $06^h\,30^m$, h_{F2}' becomes equal to h_{F2}''. This is because the wave frequency at which virtual height of F_2 region is minimum has increased to 0.834 times the penetration frequency of F_2 region. In these circumstances h_{F2}'' clearly exceeds the actual height of maximum electron density in F_2 region due to presence of F_1 region. Between $06^h\,30^m$ and $11^h\,00^m$ the wave frequency at which virtual height of F_2 region is minimum actually exceeds 0.834 times the penetration frequency of F_2 region. Consequently even the minimum virtual height of F_2 region exceeds the height of maximum electron density. During such an interval it is useless to scale h_{F2}'' unless it is corrected for presence of F_1 region. Between $11^h\,00^m$ and $16^h\,15^m$ the wave frequency at which virtual height of F_2 region is minimum is roughly equal to 0.834 times the penetration frequency of F_2 region. Consequently, h_{F2}' is approximately equal to h_{F2}'', and only one curve is shown. In this interval the common value of h_{F2}' and h_{F2}'' is necessarily greater than the height of maximum electron density, but not to the enormous extent prevailing between $06^h\,30^m$ and $11^h\,00^m$. The broken curve shows the actual height of maximum electron density in F_2 region between $05^h\,45^m$ and $17^h\,45^m$ deduced by correcting for presence of F_1 region by the method outlined at the end of Section III.

The height h_{F1}'' in Fig. 10, is appreciably greater than h_{F1}' throughout most of the daytime and gives a useful measure of the height of maximum electron density in F_1 region so far as this region has individuality (see Fig. 5). Correction of h_{F1}'' for presence of E region would decrease values by only about ten km. h_{F1}', on the other hand, does not give a useful indication of the bottom of F_1 region due to presence of E region; see Fig. 5, which refers to epoch $16^h\,15^m$ in Figs. 10 and 11.

It is worth emphasizing that minimum virtual height is ill-adapted for quantitative study of solar diurnal variation of the height of F region. At night minimum virtual height is a measure of the height of the bottom of the region. But during the day minimum virtual height usually measures the height of a level appreciably above the bottom of the region, and not uncommonly above even the level of maximum electron density. This arises partly from variable retardation in a lower region, but mainly from variation in the ratio of the wave frequency at which minimum virtual height is measured to the penetration frequency of the region under examination.

Such applications as have so far been made of the technique described in this paper indicate that it should prove a powerful method of approaching a number of ionospheric problems. Berkner, Wells, and Seaton have already used the technique in connection with an intense magnetic storm.[9] In studying regular diurnal variation of the ionosphere it is desirable to use observations averaged for a number of days. Monthly averages will shortly be available.

Throughout the conduct of this work we have had the benefit of helpful discussion with L. V. Berkner, to whom we desire to express our thanks.

[9] L. V. Berkner, H. W. Wells and S. L. Seaton, Terr. Mag. **44**, 283–311 (1939).

The Reception of Radio Echoes From Distant Ionospheric Irregularities

J. A. PIERCE AND H. R. MIMNO
Cruft Laboratory, Harvard University, Cambridge, Massachusetts

(Received September 7, 1939)

For many years it has been generally known that strong, sharply defined radio echoes occasionally return to the sending station, or to points within the normal skip-zone, after having traversed a path which greatly exceeds the round trip distance to the *F* layer of the ionosphere. Previous investigators tentatively ascribed such effects to reflection from distant mountains or from concentrations of ions in the polar regions. Studies of numerous reflection patterns strongly indicate that the delayed echoes are returned from regions where there is marked curvature of the *F* layer. A region of this sort normally occurs at the edge of the sunlit zone and can turn back a ray which may have traveled many thousands of kilometers around the dark side of the earth. Small nighttime variations in the curvature of the *F* layer are of very common occurrence and are believed to explain several phenomena, including much long period, long distance fading. These variations are frequently of cyclic character and may tentatively be attributed to wave motion in the upper atmosphere.

INTRODUCTION

CONTINUOUS records of ionospheric layer heights frequently exhibit a phenomenon which may be described as a sudden variation in the number of multiple reflections returned from the *F* layer. A record may be obtained showing, say, three multiples which have intensities decreasing with increasing path length. After a time the second and third multiples will disappear, leaving a record typical of rather "weak" reflections. This condition may persist for fifteen or twenty minutes and be followed by a period of extremely "strong" reflections, when as many as eight or ten multiples are observed. Since the coefficient of reflection at the *F* layer, for the condition of total internal reflection, is always very nearly unity, it is necessary to explain this behavior in terms of large variation in the absorption of energy in the lower layers or to assume that the reflecting surface is not (approximately) concentric with the earth.

We find many records which exhibit little variation in the strength of the echoes over periods of hours before or after the occurrence of this phenomenon. Examination of such records leaves an impression that the bursts of multiples indicate an intensity far greater than normal, while the energy returned in the intervals between bursts is below that which we would expect. This fact can be explained on the assumption that a series of waves appears at times on the surface of the *F* layer. At a moment when the surface over the transmitting (and receiving) location is concave downwards, the transmitted energy is focused upon the receiver; and when the layer, at the point of reflection, is concave upwards the received energy is subnormal. It should be noted that, if the radius of curvature is considerably greater than the equivalent height of reflection, the focusing effect will become stronger as the total path length increases, thus partially offsetting the inverse distance effect. With a very moderate assumption of curvature, in fact, we may expect the sixth or eighth multiple reflection to be returned as strongly as, say, the third.

It will be noticed that the illustration of the

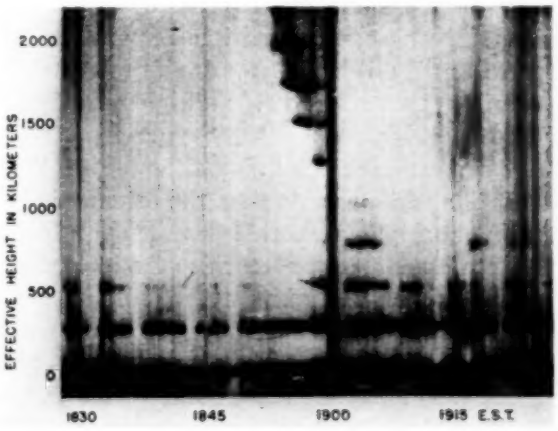

FIG. 1. "Staggered" multiple reflections from the *F* layer, recorded at 3.5 megacycles.

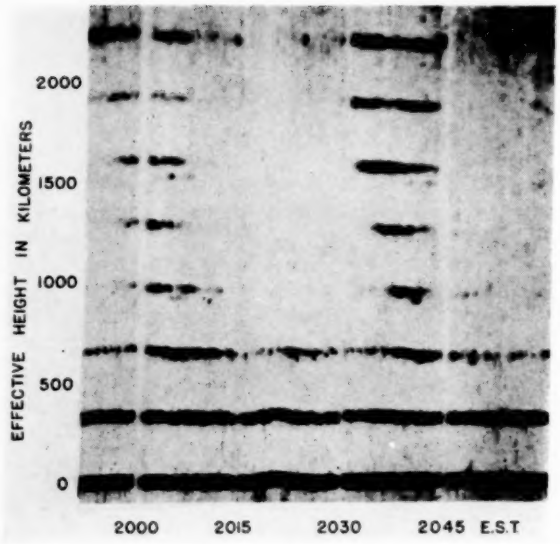

FIG. 2. Another example of "staggered" multiple reflections.

phenomenon shown in Fig. 1 exhibits a remarkable modification of the effects to be expected under the simple theory outlined above. It is characteristic, at least in the early evening, that the onset of one of these bursts of reflection is indicated by the appearance, in addition to the first two or three multiples, of two or three multiples of high order, such as the eighth and ninth. With the passage of a few minutes the highest multiples will disappear while those next below appear in sequence. Five or ten minutes after the onset of the high order reflections, we find that the entire sequence of multiple reflections has been run through in reverse order, leaving only the lowest (continuous) orders. An interesting modification of this pattern is shown in Fig. 2, where each multiple remains on the record until all have appeared, and shortly thereafter the abnormal multiples fade out simultaneously.

We can conceive of no disposition of low layer absorption which could cause the attenuation of certain multiple reflections while permitting those of higher order to return to the region near the transmitter. We must, therefore, see if our hypothesis of varying curvature of the surface of the equivalent reflector can explain the phenomenon. Unfortunately our knowledge of the geometry of such second-order curvature of the layer is indirect and inaccurate. We hope, however, to show that curvature of a type which may reasonably be expected to exist can explain records of the sort we have obtained, and to suggest possible effects of such curvatures upon communication circuits.

A Qualitative Explanation of the Phenomena

Figure 3 is a cross section of an equivalent reflector upon which we have indicated a number of ray paths. This diagram represents a condition we believe to exist and is shown in exaggerated form in order that we may trace a number of ray paths with a minimum of confusion. We have constructed this equivalent reflector by superimposing one wave-length of a cosine wave upon a plane, the cosine curve having a wave-length of 500 km and an amplitude of 5 km. Since this surface is large and at a small angle to the horizontal, we can always find a ray from the transmitter position O (which we have chosen to locate underneath a node of the curve) which is perpendicular to the reflector. This perpendicular, OA, represents the path of a ray which will be reflected directly to a receiver which shares the transmitter location. We can trace a very large number of rays which leave O at different angles to the vertical. These rays will be reflected back and forth between the earth and the layer and, in general, will not be returned to O. A few rays, however, will sooner or later strike a reflecting surface normally and will then be sent back along their outward courses and returned to O after having traversed the region between the earth and the layer a number of times. Consider, for instance, a ray which departs from O along the path OC_1, which happens in this case to be nearly vertical. This ray is reflected in sequence at C_1, C_2, C_3, and so on until it meets the earth normally at C_8. It is then returned exactly along its former path, reaching O where it can be recorded (if it be not too weak) as the eighth multiple reflection.

Figure 3 has been constructed so that we may use its nomenclature to discuss this problem qualitatively. Consider the rays which first reflect from the layer to the left of A, the point marked by the directly reflected ray. Some ray such as OB_1 will be reflected vertically downward to B_2 and then returned to O as the second

multiple. Another ray, not shown, will, at its third reflection, meet the layer normally and become the third multiple. This will continue until some ray, such as OM_1, is lost because it has not been turned back towards O before escaping from under the curved region. We find, thus, that, if we examine in sequence all rays to the left of OA, among a multitude of rays which are lost to us are a few which are returned in order as the first, second, third multiples, and so on. This is exactly what we would find in the case of a plane reflector, except that in our case the intensity of the multiple reflections is decreased because of their repeated reflection from a convex surface.

For rays reflecting originally to the right of A the situation is changed in several respects. We have found above that the ray OC_1 returns finally as the eighth multiple. Similar tracing of the rays OD_1, OE_1, and OF_1 shows that they return, respectively, as the seventh, eighth, and sixth multiples. Other rays not shown are, of course, also returned. There is a limiting angle in this case, as in the other. Rays which leave O at too great an angle to the vertical will escape at the right-hand side of the diagram before being turned back. Such a ray is shown as ON_1. In general, the farther to the right the original ray is reflected, within the limit, the sooner it is likely to return because it cannot turn back until it has reached a negative slope to the right of O. At the same time the *probability* of its being returned to O is decreasing as its original angle to the vertical increases. Two factors operate to increase the intensities of multiple reflections returned to O. In general, these rays are being reflected from a concave surface and are consequently focused back, to some extent, upon O. The other reason is that a number of rays may be returned to O after having made the same number of traverses of the region between the earth and the layer; a condition which has no parallel in the case of a plane reflector. For example, the ray paths labeled C and F, in Fig. 3, each contribute to the 8th-order multiple reflection.

We may summarize this case by stating that the normal set of multiple reflections which we would expect from a plane reflector are now represented by the reflections from the left-hand part of the diagram, these being weaker than in the plane case, especially in the higher multiples. In addition to this set of reflections we have an entirely new family reflected from the right-hand side of Fig. 3. In this family the higher multiples are more probable and there are cases of several rays combining to form a single, intense reflection of a high order, as well as a general increase in intensity due to a focusing effect. There is a limit set by the geometry of the layer which operates in such a way that no multiples of *less* than a certain order can be present in this family. If the reflections from the layer shown in Fig. 3 were to be recorded with an instrument having a finite threshold sensitivity we might well expect to see a record showing the presence of the first, second, third, sixth, seventh, and eighth multiples only.

It is obvious that the sixth multiple (made up of OF_1, F_1F_2, etc.) in this case would have a total path length much greater than six times OA, while in the actual case path lengths for the higher orders do not differ by more than one or two percent from multiples of the layer height indicated by the first reflection. This means that the layer we have sketched is indeed an exaggerated example of any similar formation which may appear in the F region. The same conclusion follows from the fact that if the layer height varied by ten kilometers from point to point and minute to minute direct evidence of the height variations would be obvious. We must, therefore, restrict ourselves to postulation of the existence of ripples in the F layer of not more than one or two kilometers amplitude and consider that the phenomena of Fig. 3 are compressed into a lateral space of not more than fifty or a hundred, rather than five hundred, kilometers. Even with this restriction, however, it seems reasonable to

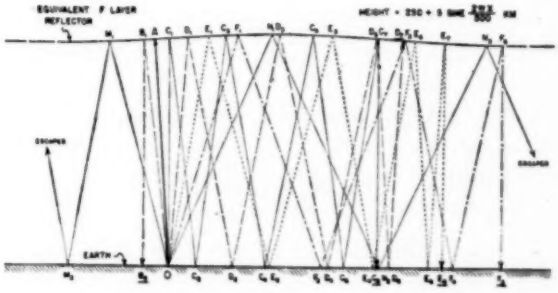

FIG. 3. A cross section of an assumed equivalent reflector showing a number of possible ray paths.

assume that this simple theory of a "dimple" in the layer can explain not only the observed variations in the number of multiple reflections, but also the fact that, from time to time, certain intermediate multiples are missing.

If we consider that a series of waves of this

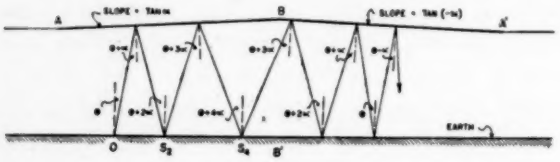

FIG. 4. A straight line approximation to the reflector of Fig. 3.

form roll across the surface of the F layer it is easy to see certain variations which are to be expected in the reflection pattern. When the transmitter is under the center of the part of the curve which is convex downwards, all the multiple reflections will be present in their normal order but with intensities which decrease more rapidly than the increase in the squares of their path lengths. When the layer is symmetrically concave towards the transmitter, all the multiples will be present but with increased intensities. At intermediate positions certain multiples may be returned with great intensities and other multiples of lower order may be forbidden or present only very weakly.

REFLECTIONS FROM A LAYER OF SYMMETRICAL TRIANGULAR FORM

This problem does not lend itself to mathematical treatment, and the complexities of a graphical solution for all possible layer curvatures and transmitter positions are insuperable. We can, however, reduce our dimple in the layer to the simple triangular form shown in Fig. 4 and draw some useful conclusions with the aid of the nomenclature we have indicated. We have given a region within which the layer has a slope of $\tan\alpha$ and, to the right, a similar zone where the slope is $\tan(-\alpha)$. From the transmitter at O, on the left-hand side of the diagram, a ray departs at an angle θ to the vertical. So long as this ray is being reflected between the part of the layer having a positive slope and the earth, its angle of incidence upon a reflector increases by α at each reflection. Similarly, when the ray is in the right-hand part of the diagram, the angle of incidence decreases by α at each reflection. Since the angle of incidence upon either the layer or the earth must be zero in order to have the ray return along its outward path it follows that θ must equal $n\alpha$ before a ray can be of interest to us, n being a positive or negative integer, or zero. This greatly limits the number of rays we must consider for any given extent and slope of the layer and any transmitter position, but it is still uncertain whether each of these possible rays will return to O, or, if so, after how many reflections.

If α is so small that we may take $\tan n\alpha = n\alpha$, it is easy to express the distance OS_2 as

$$s_2 = 2h(\theta+\alpha), \tag{1}$$

where h is the average height above ground of the points of reflection on the layer. This may be extended to:

$$s_n = n(\theta/\alpha + n/2)h\alpha, \tag{2}$$

where n is the number of times the ray has traversed the space between the earth and the layer, and θ and α are to be taken as positive for the departing ray shown in Fig. 4. This equation breaks down at the first reflection from a slope other than α, but with its aid a tabulation may be made showing the horizontal distance traversed by each of a family of rays leaving O at angles to the vertical of 0, $\pm\alpha$, $\pm 2\alpha$, etc. The distances from one point of reflection at the earth to the next of course decreases, after the rays pass under the region BA', in the same way that they had increased before.

It is easy, after these preparations, to trace each possible ray for a given set of values of h, α, the distances AB and BA', and for each of a number of transmitter positions. We find that, as O moves from A toward B, the first returning rays (as we examine those with θ positive in sequence) appear as high multiples. As O approaches B', the energy returning to O is contained in lower and lower multiples until, when O is at B', all multiples are present as though reflection were taking place from a plane surface. As O moves from B' toward the right, the procedure repeats itself in the reverse direction.

We have traced the rays corresponding to the case of a dimple 2 km deep and 200 km in total

length in a layer at a height of 250 km. The results are exhibited in Fig. 5 where the solid lines represent reflections of rays which departed with θ positive, and the dotted lines the reflections for θ negative. Widened lines indicate cases where two or more rays return to make up a given multiple reflection.

A number of complex rays return as even higher multiples than those shown. These are the ones which depart with θ negative, reflect first at the left of O, then at the right, and finally return to O after very long delays. They are omitted from the figure because their long path length probably reduces their intensity to a value which we could not detect.

The Effects of Asymmetry

Figure 5 represents the reflection pattern we would expect to record if a triangular wave of this sort rolled over the transmitter. It differs from the most usual type of record chiefly in its symmetry. Occasionally, in the middle of the night, a pattern of this sort is observed, one being shown in rudimentary form at the left of Fig. 6. In general, however, the phenomenon is characteristic of the three or four hours after sunset, and at this time the right-hand part of Fig. 5 is absent. The most reasonable explanation of this asymmetry is that waves are superimposed upon a sloping layer. Examination of a number of records gives us an average increase in layer height (from west towards the east) of about one kilometer in a hundred, during the evening. If we superimpose the ripple of Fig. 4 upon this slope we find that the slope AB is only one-third of that between B and A'. The region BA' is now very effective in turning back rays which originated at the left of B, but rays which are traveling towards the left will, in general, escape. This asymmetry is shown in Fig. 7, which was drawn up under the same conditions as Fig. 5, except for the slope of -0.01 upon which the ripple was superimposed. The seventh, eighth, and ninth multiples appearing just to the right of O are part of a sort of second-order pattern most of which is contained in multiples higher than those shown. This part of the pattern is sensitive to small changes in the geometry of the layer and may be expected to be rudimentary in many cases. It is probable that, when it is present, it is blurred sufficiently to merge it with the primary pattern, the combination giving a record of the type of Fig. 2.

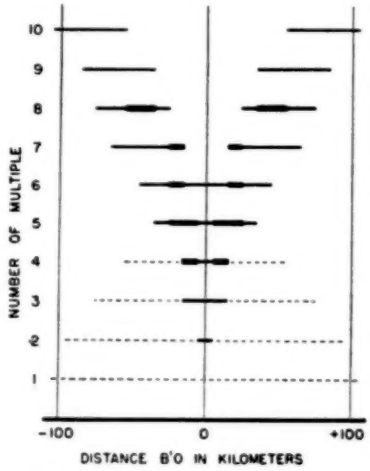

Fig. 5. A computed reflection pattern. The ordinate may be expressed in time to make this figure comparable with the experimental records.

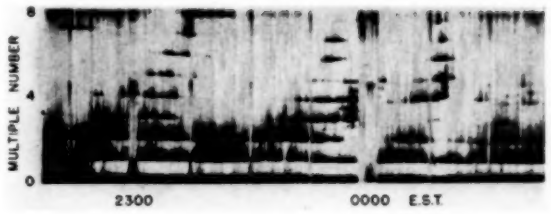

Fig. 6. An example of the variation of multiple reflections with time. The portion of the record centered at 2300 should be compared with Fig. 5.

The Magnitude of Ripples in the *F* Layer

We have remarked that the amplitudes of ripples of this sort cannot be more than one or two kilometers, or the height variations would be directly measurable. Since this is the case, we expect that the phenomenon of the forbidden multiple reflections should be observed only when the receiver and transmitter are very close together. This is because some of the transmitted rays must pass laterally beyond the receiver and return to it, and this is impossible unless θ and α, in Fig. 4, are of the same order of magnitude. For transmission over distances comparable to the

height of reflection, θ must be many times larger than any angle between the earth and the ionized layer could possibly be. This conclusion has been checked experimentally. The records reproduced in Figs. 1, 2, and 6 were made with the receiver within 250 meters of the transmitter, while simultaneous records were made at a distance of 240 kilometers and showed no indication of the phenomenon. As a check on the maximum slopes in the layer, the receiving equipment was removed to a location about four kilometers from the transmitter and operated there for about ten days. The records made at this distance exhibited the effects of *focusing* in the variation of the number of multiples, but the echelon formation of multiple reflections was present only in rudimentary form in one or two cases although the records made at 250 meters distance immediately before and after the removal indicated a number of occurrences nearly every day.

If we were to draw a circle through the points A, B, and A' in Fig. 4 we would find a radius of curvature of 2500 km, and this figure may well be increased to 4000 km to allow for the defocusing effect of the convex surface of the earth. It is

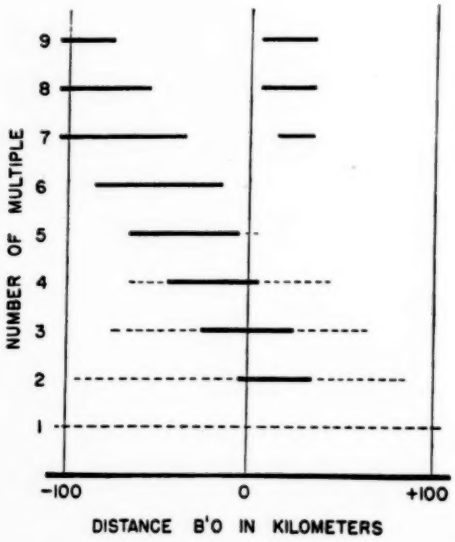

FIG. 7. The computed reflection pattern corresponding to the layer of Fig. 4 and the records of Figs. 1 and 2.

easy to calculate the variations in intensity of a ray reflected between a cylindrical mirror and a plane reflector, as the number of multiple reflections increases. The results of such a calcu-

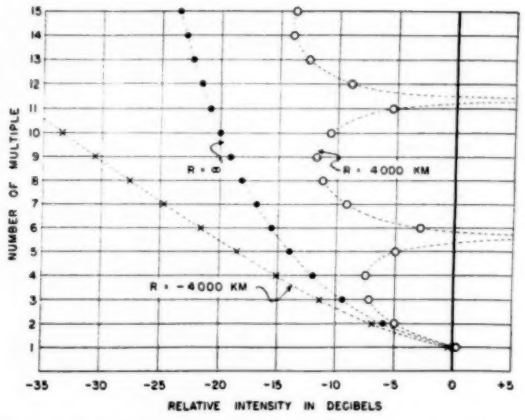

FIG. 8. The computed variation in intensity of a ray reflecting between a plane and a cylindrical mirror separated by 300 km.

lation, for a layer height of 300 km and a radius of curvature of 4000 km, are shown in Fig. 8. Perfect reflection and no attenuation in the absorbing layers have been assumed, but the effect of attenuation may be simulated by giving the ordinates in the figure an arbitrary negative slope. Since a truly cylindrical curvature cannot be expected, it is probable that the cusps computed for the concave case are blurred out of existence so that the values fall more or less along a smooth curve which, for multiples beyond the second or third, varies inversely as the multiple number rather than as the square of that number. The most interesting conclusion from the figure is that we may expect the high multiples, such as the eighth or tenth, to vary as much as 20 db if the reflecting surface becomes alternately concave

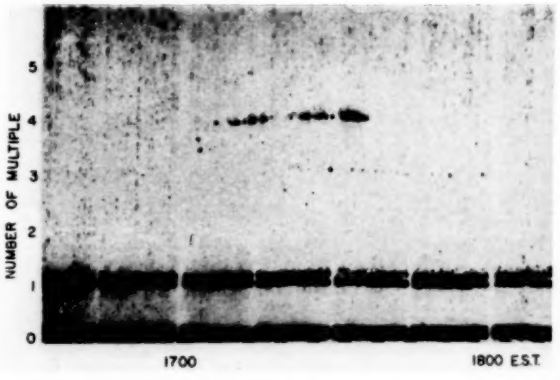

FIG. 9. A record showing the presence of a fourth multiple reflection which is stronger than either the second or the third.

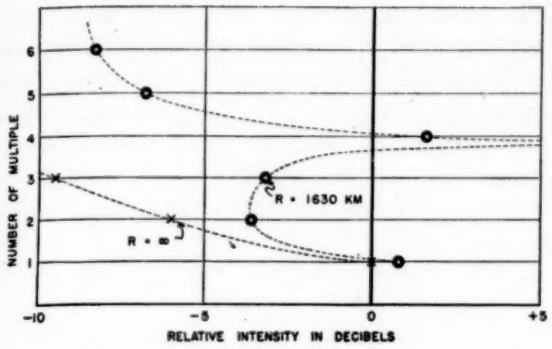

FIG. 10. The computed variation in relative intensity of several multiple reflections for a case where the curvature is large.

and convex. An effect of this magnitude would be caused by the passage of a sine wave, such as that shown in Fig. 3, having a wave-length of 200 km and an amplitude of 0.25 km.

Occasionally random curvatures of greater magnitude occur. The record corresponding to one of these is reproduced in Fig. 9 and, although made with the receiver sensitivity below normal, shows that the fourth multiple was, for a time, nearly as strong as the first, and that the second is absent while the third is represented by a very weak trace. The variations in intensity compare favorably with those computed for a cylindrical curvature having a radius of 1630 km which are exhibited in Fig. 10.

THE FADING OF LONG DISTANCE SIGNALS

It is obvious that the parts of this discussion dealing with phenomena which are detectable only in the special case of vertical incidence observations can have little importance in practical communication. They have, however, enabled us to postulate the frequent occurrence of significant changes in the curvature of the *F* layer. For communication over short distances these changes are unimportant, but when the total length of the transmission path approaches the radius of curvature they become very interesting. It is possible that much of the long period fading observed on long distance circuits may be a result of this effect. This type of fading occurs too slowly to be a result of polarization changes or interference between a number of rays. It seems to us that the hypothesis of varying attenuation in the absorbing strata may not explain the cases of slow fading observed at night and at high frequencies, while the mechanism we are suggesting is not, within limits, a function of frequency. An example of what may well be a focusing effect is shown in Fig. 11, which is a record of the field strength of a short wave broadcasting station operating on a frequency of about 12 megacycles at a distance from Cambridge of 5200 kilometers. The transmission begins at 0115 E.S.T., when the ionization density has fallen so low that only one ray path is effective. This is indicated by the absence of the rapid fading caused by the interference between several rays. The signal remains steady and pleasant to hear but with a smooth fading cycle

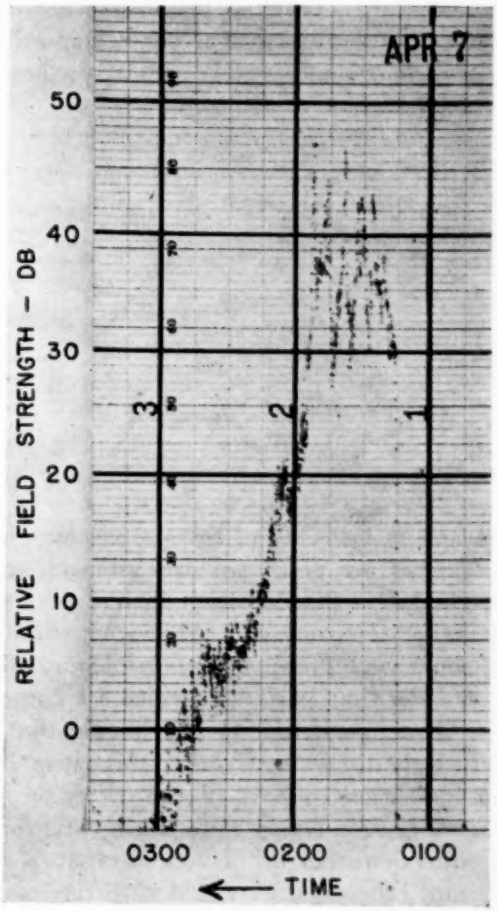

FIG. 11. A record of the field strength of a transatlantic short wave broadcasting station. Note that the time scale reads from right to left.

having a period of about ten minutes and a full amplitude of nearly 20 db. This lasts until about 0155 E.S.T. when the ionization density falls below the limiting value, the signal strength decreasing by 30 db in the next few minutes. It appears to us that this slow, cyclic fading without change in the general signal level is evidence of the focusing effect we have postulated.

An Hypothesis of Ripple Formation

Under the conditions obtaining in the *F* region in the evening of a magnetically quiet day, the equivalent height of reflection increases by about twelve or fourteen kilometers per megacycle in the region near three or four megacycles. This can be interpreted to mean that the density of ionization increases by about five percent for one kilometer increase in height. If, then, irregularities in height of this order of magnitude are to be explained by assuming that some non-uniform source of ionizing energy is operative, the *local*

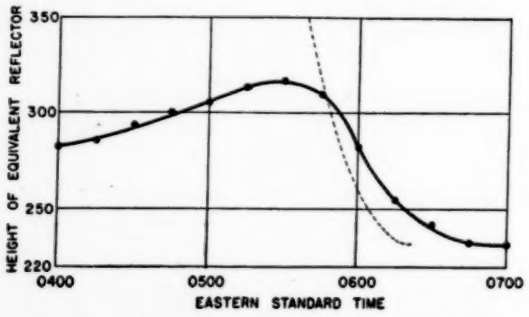

Fig. 12. The variation of *F* layer height with time in the sunrise period (3.5 megacycles; November, 1938).

variations in intensity of this source must be of the order of ten percent of the intensity which would maintain the ionization in equilibrium at a value of 10^5 electrons/cm^3. While such variations in incident energy are not impossible and are, in fact, smaller than variations which are known to affect the *E* region, it is improbable that the effect we are discussing is due to this cause. If the total ionization in part of the *F* region were changing in such a way as to cause variations in the equivalent height of reflection at a fixed frequency, the changes would give rise to phenomena of other types, such as cyclic variations in the critical frequency, which have not been reported.

We prefer to believe, therefore, that these effects are due to a temporary redistribution of the free electrons already existing in the lower part of the *F* region. The simplest mechanism which will cause such a redistribution is an atmospheric compressional wave. We have noted that our phenomenon occurs generally in the period immediately following sunset. This permits us to postulate a disturbance, set up by the cooling of the upper atmosphere at sundown, which is propagated in the direction opposite to that of the advancing shadow. It will be noted that this direction of propagation, together with the slope in the equivalent layer, is necessary to explain the observed asymmetry in our records.

It is difficult to estimate the velocity of propagation in the tenuous medium which exists at the height of the *F* region, where the mean free path is probably several kilometers and the temperature and pressure can only be guessed at, but it is obvious that waves can be set up whose length is large compared to the mean free path. Such waves would cause compressions and rarefactions of the free electrons in the *F* region in essentially the same ratio as in the atoms or molecules of gas if, as is unquestionably the case, the period of half-recombination is long compared to the period of the wave, and these changes in electron density would appear as changes in the height of the equivalent reflector.

We are not able to deduce the wave-lengths of these disturbances accurately, but they seem to be of the order of two or three hundred kilometers. The records which exhibit a definite cyclic character give a mean period of about 20 minutes, although it often varies between half and twice this figure. The velocity corresponding to these estimates is some six or eight hundred kilometers per hour or, say, 200 meters/second. This figure is of the magnitude of the velocity of sound and is, therefore, somewhat interesting. It is barely possible that a careful study of this phenomenon, involving simultaneous observations at two or more locations, might lead to a reasonably accurate figure for the velocity of sound at the 300-kilometer level.

Focusing of Echoes at Sunrise

In contemplating the effects of changes in curvature of the equivalent *F* layer reflector it is

necessary to consider the curvatures associated with sunrise. This region is important because there is a large change in effective height in a horizontal distance of a few hundred kilometers, with correspondingly large curvatures, both convex and concave. Fig. 12 shows the average variation of equivalent height with time for a number of magnetically quiet days in November, 1938. The dotted line indicates the height at which the sun's rays strike the region after grazing an "opaque" sphere of 6600 km radius, and is shown here only because it is interesting to note that the layer height begins to fall before there is any directly incident ultraviolet radiation, indicating that diffusion is an important factor at these heights.

If we assume that there is no variation in electronic density with longitude (as distinguished from time) the heights of Fig. 12 may be plotted as a function of distance by use of the fact that sunrise, in these latitudes, advances at some 1200 km/hour. On this sort of graph radii of curvature may be measured which vary, in this average case, from infinity to about 5000 kilometers. If the curvature is somewhat greater than average we should immediately expect an increase in intensity of the echoes returned from the concave region exactly as in the case of the random curvatures discussed above. Under magnetically quiet conditions and when the frequency of observation is well below the critical frequency, this deduction is confirmed experimentally. An example of this sort of record is exhibited in Fig. 13, where the focusing of the fourth and fifth multiple reflections is clearly marked.

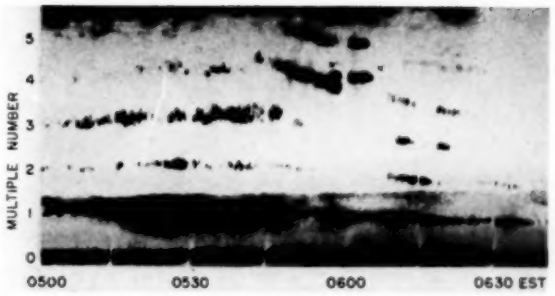

FIG. 13. A record showing the focusing of the fourth and fifth multiple reflections at sunrise (November, 1938).

THE RETURN OF LONG DELAY ECHOES

At this point we should emphasize the distinction between the sort of concavity in a layer which can cause the focusing of a beam of incident energy, and the effect of slope in a layer which can turn back a ray which has had a component of its path in the direction of the lower equivalent height. The latter effect is of especial importance in the sunrise zone. A ray which has been propagated along the dark side of the earth will, upon entering the sunrise zone, meet the reflecting layer and the earth at an angle of incidence which decreases with each reflection until it reaches zero, or until the ray passes on into a region in the sunlit zone where the layer is again concentric with the earth. If the angle of incidence is reduced to zero and if the horizontal component of the direction of the ray is perpendicular to the edge of the shadow, the ray will be returned toward the transmitter along a path similar to that followed on its way to the sunlit zone. This mechanism can be used to explain echo paths of long delay, such as some which have been ascribed to "*G*" layers.

In Fig. 14 we have shown the average *F* layer curvature in the sunrise zone deduced from Fig. 12, with the addition of a sporadic *E* layer. Between these layers we have indicated a ray which enters the region from the left and which finally meets a reflecting layer perpendicularly at

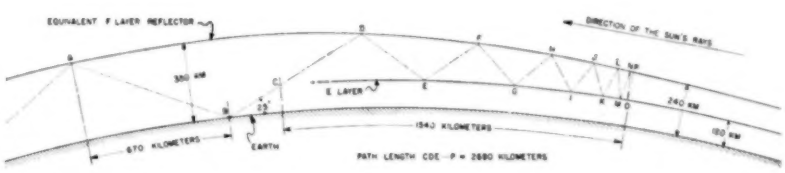

FIG. 14. A cross section of the ionosphere in the sunrise zone, showing the path of a ray which may be returned to the neighborhood of the transmitter.

P so that it will return along its outgoing path. This construction is intended to illustrate the mechanism of reflection of one echo of long delay which was observed, using a sweep-frequency technique, on November 7, 1938. This echo was delayed by almost exactly 0.1 second, corresponding to a total equivalent path of 30,000 km. The delay was not, within limits, a function of frequency, indicating that an explanation of its existence must be based upon the geometry of the reflecting surface rather than upon variations in ionic density. At the time of this record, which is not clear enough for reproduction here, the great circle distance from Cambridge to the point where a ray would meet the sunrise line perpendicularly was somewhat less than 12,000 kilometers. Thus, allowing for the vertical components of the ray path shown in Fig. 14, the reasonable assumption that the average velocity of propagation was slightly less than the velocity of light would enable us to explain this echo completely. The E layer ionization was assumed because the record showed a tendency towards a family of reflections differing in their "apparent height" by about fifty kilometers, which is approximately half the spacing between the two layers at the point where reflection occurred perpendicularly. A second ray, after following a path almost identical with that shown in Fig. 14, could be turned back (by making two successive reflections at the same angle of incidence) without having encountered either layer normally but having traversed the region between the layers once more (or less) than the ray shown in the figure. Thus a family of reflections having very long delays which differ by multiples of the time corresponding to the distance between the E and F layers could be returned to the transmitter. The (sporadic) E region ionization may have existed only in the sunrise zone, or may have extended a considerable distance along the path towards the transmitter. The latter assumption would help to explain the fact that the reflections were strong enough to be recorded, because propagation between the E and F regions should be subject to much less attenuation than propagation between the F region and the earth, where the ray would pass repeatedly through low level absorbing strata.

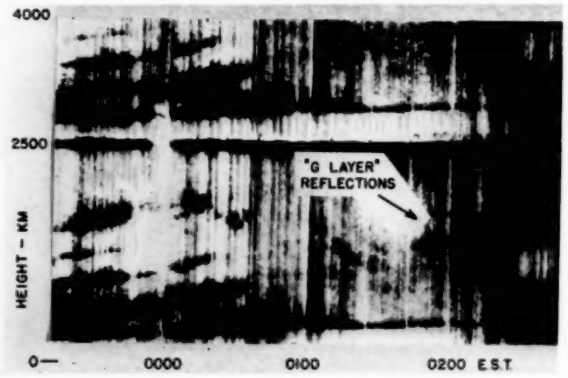

FIG. 15. A record showing, at the right, a group of G layer reflections. Note that the base line is below the bottom edge of the photograph, the lowest continuous trace being the first F layer reflection. Part of the pattern is repeated and normally the height scale would be repeated also.

Figure 15 is a reproduction of a record exhibiting an unusually uniform structure of "G layer" reflections which appeared at about 0200 E.S.T. at the end of April, 1934. Here the apparent "height" of the group is about 1500 km. Since the use of 60-cycle pulses causes the whole reflection pattern to be repeated with a period corresponding to 2500 km of effective height, the equivalent distance is 1500 plus a multiple of 2500 km. At this time the edge of the sunlit zone was about 3000 km north by east of Cambridge. Comparison of this pattern with Fig. 14 will indicate that the effective distance of the reflections was about 4000 km, the additional 1000 km being accounted for chiefly by the vertical components of the ray direction. The family of reflections referred to above is seen distinctly in this record, and their spacing is that to be expected if the height of reflection, at the point where the rays turned back, was about 220 km.

In conclusion, we should refer to a paper[1] which came to our attention after our oral presentation of this report. The authors of the paper concluded an attempt to explain certain complex echo patterns with the paragraph:

"However, we have to consider the other possibility put forward by one of us, that at times the ionized regions may be an undulatory

[1] G. R. Toshniwal, B. D. Pant, R. R. Bajpai and B. K. Verma, Proc. Nat. Acad. Sci. (India) **6**, 161–174 (1936).

surface with various ledges here and there. The presence of such a structure along with the possibility of arrival of echoes from angles other than the vertical can easily account for the complex echoes. If this view of the production of the complex echoes is accepted, we have to search for the agency which can cause such a structure of the ionosphere."

JANUARY 15, 1940 PHYSICAL REVIEW VOLUME 57

Nuclear Isomerism in Selenium and Krypton

ALEXANDER LANGSDORF, JR. AND EMILIO SEGRÈ
Radiation Laboratory, Department of Physics University of California, Berkeley, California

(Received November 4, 1939)

A study has been made of the radioactive chain

$$Se^{83} \rightarrow Br^{83} \rightarrow Kr^{*83}.$$

The last member is an isomer of the stable krypton of mass 83. An isomeric pair $Se^{79 \text{ or } 81}$ has been found and the two genetically related nuclear isomers have been separated chemically. Kr^{83} and Br^{83} are produced by the fission of uranium and thorium as well as by the bombardment of selenium with deuterons and slow neutrons.

IN the course of an investigation of radioactive products of thorium and uranium fission, we observed a radioactive krypton growing out of a radioactive bromine with a period of 2.4 hours. This bromine period coincides with that assigned by Snell[1] to Br^{83}, which he found to be the product of the decay of Se^{83} produced by deuteron bombardment of selenium. He did not, however, observe radioactive krypton growing out of the bromine. We therefore repeated this deuteron bombardment of selenium and were able to extract from the products a radioactive isotope of krypton having the same period as the fission product, and emitting the same type of radiation. Since the radioactivities of selenium and bromine appeared rather complicated, we thought it desirable to go over the different points in detail.

The main results of our investigation are the identification of an excited state of the stable Kr^{83}, the identification and chemical separation of an isomeric pair in $Se^{79 \text{ or } 81}$, and confirmation of the identification of Se^{83}. Although discovered by Snell, there had been some uncertainty as to the true half-life of Se^{83}.

[1] A. H. Snell, Phys. Rev. **52**, 1007 (1937).

THE CHAIN $Se^{83} \rightarrow Br^{83} \rightarrow Kr^{*83} \rightarrow Kr^{83}$

We have produced Se^{83} by means of slow neutrons with the reaction Se^{82} (n,γ) Se^{83} and by deuterons with the reaction Se^{82} (d,p) Se^{83}. The latter reaction of course gives much more concentrated samples and has been used for all the chemical separations. In order to determine the period of Se^{83}, repeated extractions of Br^{83} are necessary to identify its parent in the complex of selenium activities. The Br^{83} can be recognized easily from the fact that it is the only bromine isotope giving rise to an active krypton.

In order to make repeated extractions of bromine, the bombarded selenium was converted to selenic acid free of the primary 34-hour bromine activity produced by the reaction Se^{82} $(d,2n)$ Br^{82}. Preliminary experiments with the more rapidly prepared selenious acid had shown serious difficulties as a result of the apparent impossibility of achieving a complete extraction of bromine from this material, so that the longer chemical preparation of selenic acid was found to be necessary. To the selenic acid solution we added periodically five milligrams each of potassium bromide and iodide, and precipitated them with silver sulphate solution. Decay curves of the bromine activities thus

extracted are shown in Fig. 1, curves I–VII. Curves I–V all indicate a half-life of 140 minutes which is assigned to Br^{83}. In each of these curves the point marked with a cross represents the bromine activity extrapolated back to the time of extraction. Curve VIII, drawn through these crossed points, indicates that the period of the parent Se^{83} is about 30 minutes.

The absorption curve of the Br^{83} beta-rays gives an end-point of 1.05 Mev by Feather's method.[2]

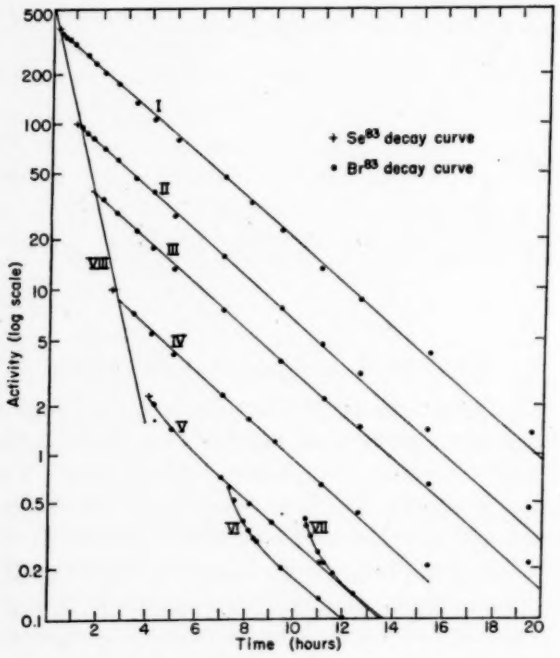

FIG. 1. Decay curves of successive bromine extracts from deuteron-bombarded selenium. The crosses represent the bromine activity extrapolated back to the time of extraction. The curve VIII through the crosses represents the decay curve of the parent selenium 83.

Decay curves of the last two extracts of bromine from the active selenic acid are shown in curves VI and VII of Fig. 1. They are distinctly anomalous in intensity and slope. Although we have not investigated this anomaly, it might be due to a short-lived bromine activity which grows out of a selenium activity with a period longer than that of Se^{83}. Since the curves of Fig. 1 are compounded of the decay curves of Br^{83} and another unknown activity, there is some uncertainty in the assignments of

[2] N. Feather, Proc. Camb. Phil. Soc. **34**, 599 (1938).

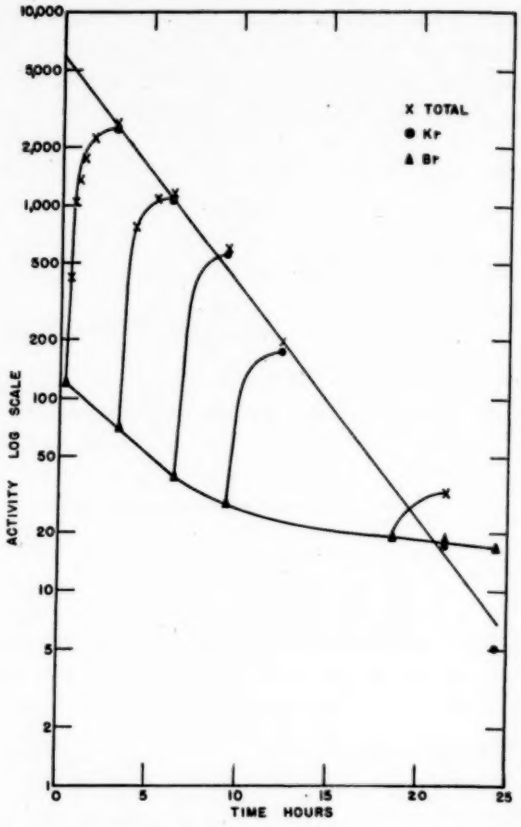

FIG. 2. Activity of an emanating bromine sample inside an ionization chamber. Triangles represent activity due to bromine; crosses represent total activity. The growth of activity is all due to accumulation of krypton. The dots give the activity due to krypton alone after 3.5 hours of growth. The line through the dots gives the period of the bromine parent of the krypton activity.

the periods of Se^{82} and Br^{83} from the data of this figure. We have, therefore, made use of the simple activity of Kr^{83} to confirm the above assignments of periods to the Br^{83} and Se^{83}.

Radioactive krypton can be extracted from a solution containing Br^{83} simply by boiling the solution. However, rather prolonged boiling combined with bubbling of a gas stream through the solution is required for a quantitative extraction. Although we have used this technique, we found it more convenient to use emanating solid samples when we wanted to make repeated extractions or when krypton was wanted with but little air admixture. We will describe later how to prepare emanating samples.

The decay curve of krypton is a perfect exponential; we followed it over a factor of ten

thousand, and the half-value period is 113 minutes.

The relation of Kr^{83} to Br^{83} has been shown by the following experiment. An emanating sample of radioactive bromine was suspended inside an ionization chamber which measured the sum of the bromine and krypton activities. The air was pumped out and replaced by fresh air at equal time intervals, so that each exhaustion was an extraction of krypton. The decrease of activity due to this extraction is, therefore, a direct measure of the krypton activity, and the residual activity immediately after letting in fresh air is bromine activity. In Fig. 2 the crosses represent the total activity, the triangles show the bromine activity, and the circles represent the krypton activity. Since the times of extraction of krypton were uniformly spaced, the curve of krypton activity is simply the decay curve of its parent substance. It will be seen that the period of the bromine so obtained agrees with that of the Br^{83} found directly.

As indicated before, we found it desirable to check the period of the Se^{83} by the Kr^{83} activity because we were not sure that the bromine decay curve was simple. This was done by extracting bromine periodically from active selenic acid with carbon tetrachloride. Half of the extracted bromine was converted into an emanating sample and half was measured directly. All the emanating samples were allowed to accumulate krypton for 3.5 hours so that the activity due to krypton after this interval of time was proportional to the amount of parent bromine initially present. Fig. 3 shows the results of measurements of the krypton and bromine activities. The initial bromine and krypton activities for different extractions are in a constant ratio. This shows that practically all of the bromine activity in our extracts is due to Br^{83} and affords also an independent check on the half-life of Se^{83}.

Krypton 83 emits a particularly soft radiation. We have been able to detect it only by introducing the gas into an ionization chamber. In order to have an estimate of the energy of the electrons emitted, the active krypton was admitted into a spherical ionization chamber of 12.6 cm diameter, which had previously been evacuated. Air was then admitted and Fig. 4 shows the ionization as a function of pressure. Similar curves have been obtained in x-ray work, and a discussion of their shape as a function of the absorption coefficient of the electrons is to be found in Bothe's article.[3] His discussion refers to a plane condenser, whereas our experimental arrangement corresponded to a spherical condenser, which introduces some uncertainty in the determination of the dimension of the vessel to be considered as corresponding to the plate

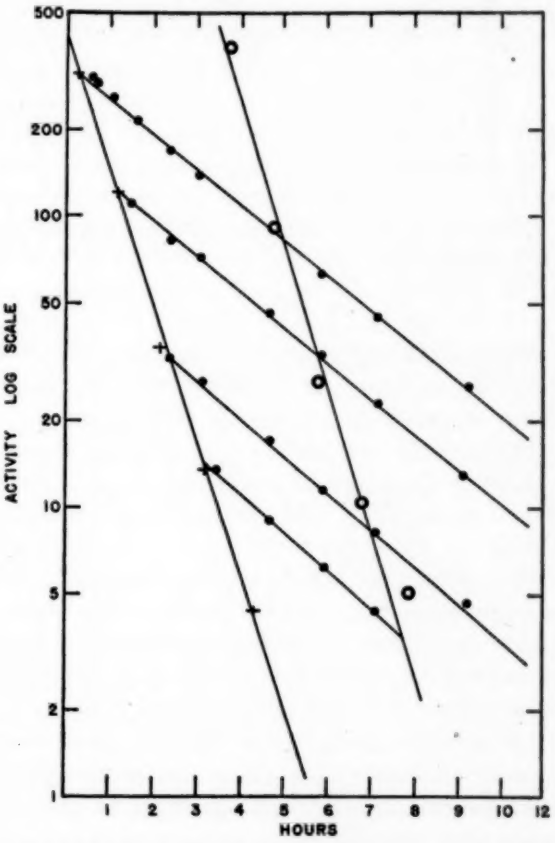

FIG. 3. Successive extractions of bromine from selenium. Dots and crosses same as in Fig. 1. Circles represent krypton activity which grows out of bromine extracts in 3.5 hours.

separation of the plane condenser. Nevertheless, the absorption coefficient varies so rapidly in the region concerned that a considerable error in the estimation of the absorption coefficient makes only a small error in the estimation of the energy. From Fig. 4 the absorption coefficient

[3] W. Bothe, Geiger, Scheel, *Handbuch der Physik* (Berlin, 1933), **23/2**, 26.

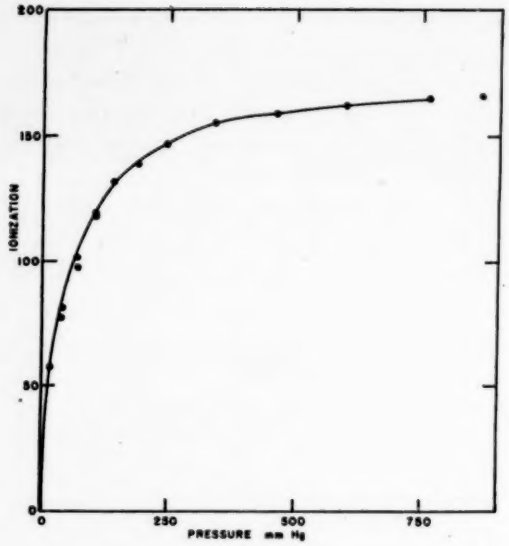

FIG. 4. Ionization due to krypton electrons as a function of the air pressure in the spherical ionization chamber of diameter 12.6 cm.

is about 1630 cm²/g, and the practical range is about 4.2 mg/cm². Both these results give an energy of about 35 kev for the electrons.

The natural interpretation of the observations is that the electrons are conversion electrons of a 49-kev gamma-ray. This energy is obtained by adding to the observed electron energy the binding energy of the *K* level in krypton. This interpretation is further confirmed by the detection of the characteristic x-radiation of krypton by absorption measurements on a gas sample. X-rays were detected with a "freon" chamber. We found that arsenic absorbs the x-rays more strongly than selenium. Since the energy of the characteristic *K* radiation of krypton lies between the absorption edges of arsenic and selenium, this experiment shows conclusively that the x-rays are characteristic of krypton.

Since there is no doubt that the parent of this radioactive krypton is the bromine isotope of atomic weight 83 (i.e., Br^{83}), we must conclude that this krypton is a metastable excited isomer of stable Kr^{83}.

In this connection, we want to point out that Kr^{83} is known to have a nuclear spin of 9/2.[4] According to the computations of H. Hebb and G. E. Uhlenbeck,[5] the observed energy and half-life indicates a change of spin of four, so that the excited state probably has a spin of $\frac{1}{2}$.

We think it is within the possibilities of the cyclotron to prepare enough excited Kr^{83} to observe the hyperfine structure in its spectrum and its change when Kr^{83} reverts to the ground state. This would permit a direct test of the Weizsäcker theory of nuclear isomerism.

SELENIUM 79 OR 81

The selenium activity is so complex that we thought it worth while to look for isomers, as a means of furthering the analysis, especially since we know that the chemical method of separation of nuclear isomers[6] is applicable in the case of tellurium[7] which is chemically similar to selenium.

Radioactive selenic acid and hydrochloric acid were mixed so that the final solution was four

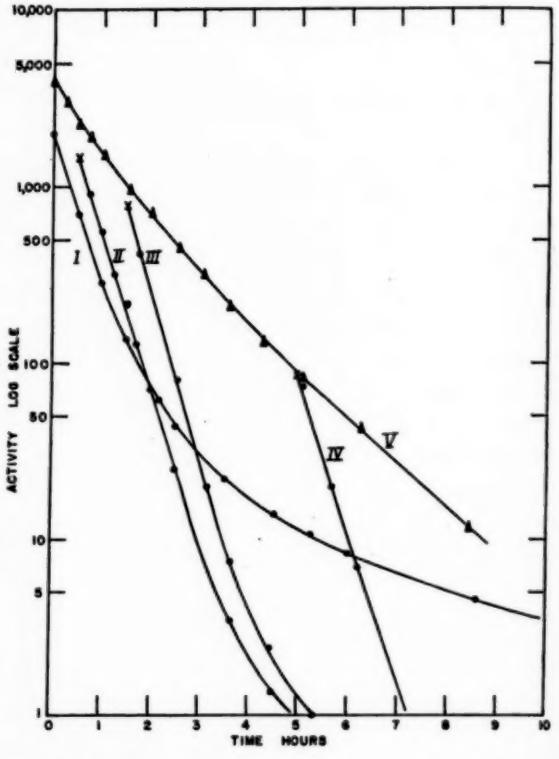

FIG. 5. Successive extractions of the lower isomeric state of $Se^{79, 81}$ from the total selenium activity. Dots for separated isomer activities; crosses for the initial activity give the period of the parent isomeric state approximately. Total selenium activity (triangles) is on a different scale.

[4] H. Korsching, Zeits. f. Physik **109**, 349 (1938).
[5] H. Hebb, G. E. Uhlenbeck, Physica **5**, 605 (1938).
[6] E. Segrè, R. S. Halford and G. T. Seaborg, Phys. Rev. **55**, 321 (1938).
[7] G. T. Seaborg, J. J. Livingood and J. W. Kennedy, Phys. Rev. **55**, 794 (1939).

normal with respect to hydrochloric acid. The solution was then saturated with sulphur dioxide. Under such conditions selenic acid is reduced very slowly whereas selenious acid is reduced rapidly to selenium which can be filtered out of the solution. To this selenic acid solution we added at intervals of time selenious acid and separated the elementary selenium. Fig. 5 shows a decay curve of the parent selenium and of the extracts. The extracts show a half-life of 19 minutes, whereas the parent selenium has a half-life very close to one hour.

The assignment of this isomeric pair to $Se^{79 \text{ or } 81}$ is based upon the observation that it can be obtained also by neutron bombardment of bromine, which has only two stable isotopes, Br^{79} and Br^{81} in equal amounts. The activity of selenium obtained from bromine by the (n,p) reaction is much less complex than the activity which one obtains by deuteron bombardment of selenium. Using this latter method of preparation of $Se^{79 \text{ or } 81}$, we could directly measure the half-life of the upper isomeric state with good accuracy. It is very close to one hour, in agreement with an observation of Snell.[1]

We also checked that this isomeric pair of selenium activities does not decay into any active product, a result that is to be expected.

These observations explain the discrepancy between the results of Snell,[1] and Bothe and Gentner.[8] The latter probably observed the lower isomeric state of $Se^{81 \text{ or } 79}$ which they obtained by gamma-ray bombardment; Snell, on the other hand, had not distinguished Se^{83} from the lower isomer of Se^{79-81}.

By the chemical method we extracted the 20-minute lower isomeric state of $Se^{79, 81}$ produced by bombardment of bromine and found thereby that we had extracted at least one-half of the beta-ray activity of the parent selenium. It must, however, be borne in mind that the yield of our separation, due to its chemical features, may be rather low. It is quite possible, therefore, that all of the beta-ray activity of $Se^{79, 81}$ is due to the lower isomeric state. The absorption curve in aluminum of the beta-rays of both activities is the same and gives by Feather's method[2] an upper limit of the energy of about

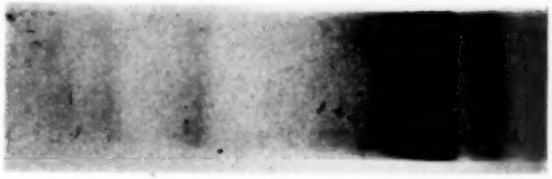

FIG. 6. Beta-spectrogram of Se bombarded with deuterons showing K- and L-conversion electrons ejected in the isomeric transitions of $Se^{79 \text{ or } 81}$ and Kr^{83}.

1.5 Mev. These results indicate that the upper state decays to the lower one, which in turn disintegrates to bromine. The conversion electrons which correspond to the transition from the upper to the lower state have been observed as a very soft component in the absorption curve of the electrons emitted by the one-hour activity and have been detected in a magnetic spectrogram. Both K- and L-conversion lines, are illustrated in Fig. 6. The energy of the K-conversion electrons is about 85 kev, corresponding to an energy difference between the isomeric states of 98 kev. The same value has been found by absorption measurements on the electrons.[9]

CHEMICAL PREPARATIONS

Selenic acid

The bombarded selenium was dissolved in nitric acid containing some bromide ion, then evaporated to dryness to eliminate most of the strong primary bromine activities, redissolved and boiled with sodium hydroxide, then filtered to remove metallic contaminations such as copper from the target plate. The solution was acidified with nitric acid and treated with an excess of normal silver nitrate to precipitate silver selenite. This compound was washed, suspended in water, and the selenite was oxidized to selenate by treatment with bromine, of which the excess was removed by boiling. The filtered solution contained selenic acid free of primary radioactive bromine.

Separation of bromine

As an alternative method to the precipitation of silver bromide out of selenic acid solution as already described, we have added equivalent

[8] W. Bothe and W. Gentner, Zeits. f. Physik **112**, 45 (1939).

[9] We thank Mr. A. C. Helmholz for photographing the line spectrum of electrons.

quantities of sodium bromide and sodium bromate in order to liberate bromine, which was then extracted with carbon tetrachloride. This method, while not so efficient as the former precipitation technique, has the advantage of giving the bromine in a form well suited to the preparation of emanating samples.

Emanating samples

We prepared silica gel impregnated with silver nitrate by immersing the granular silica gel (8 to 16 mesh) in 0.1 *N* silver nitrate solution for about 15 minutes. The gel was then drained and dried for several hours on a hot-plate at about 70°C. The radioactive bromine was prepared in the form of free bromine with several milligrams of bromine carrier, dissolved in about 5 cc of carbon tetrachloride (free of any visible water layer but not specially dried). Several cubic centimeters of the silver nitrate-impregnated silica gel were added to this solution and shaken with it until the bromine color had completely disappeared. This usually took about 15 minutes. The silica gel was drained and dried on the hot-plate for about one-half hour. It was then ready to be placed in a vessel for collection of the emanating krypton. This method of preparing an emanating solid has been found to be equally applicable in the case of radioactive xenons which are produced by decay of iodine extracted from thorium and uranium fission products.

Other observations

The preceding work is not a complete investigation of the radioactivities produced by bombardments of selenium. Our work indicates there are several points requiring further study. Perhaps there is a radioactive bromine with a half-life less than one-half hour which grows from a radioactive selenium with a half-life considerably longer than that of Se^{83}. This unidentified activity has already been discussed in connection with Fig. 1, curves VI and VII.

We have also observed an activity in selenium bombarded with deuterons with a half-life of more than a day. Since this longer-lived activity tends to be selectively extracted from the parent selenate solution when making the first isomeric separation, for $Se^{79 \text{ or } 81}$, it may be that this longer-lived activity is itself the lower isomeric state of another as yet unknown selenium. Fig. 5 shows evidence on which these conclusions are based. Curve I, Fig. 5, shows a strong long-lived tail in the first isomeric extract. This tail is much less pronounced in later extracts. It cannot be due to an imperfect isomeric separation, because the tail actually shows a longer period than is shown by the *total* selenium activity in curve V of Fig. 5. However, the possibilities of the presence of impurities in the selenium and of the formation of arsenic during the bombardment have to be borne in mind in appraising the results reported in this paragraph.

Since Br^{83} and Kr^{83} are observed in the fission of thorium and uranium, it is probable that Se^{83} is also a product of the fission. It has not, however, been observed directly. An extract of selenium from neutron irradiated uranium showed the presence of several different activities, which prevented a simple confirmation of the presence of the Se^{83} activity.

Our thanks are due to Professor E. O. Lawrence for his interest in this work. The financial support of the Research Corporation and of the Rockefeller Foundation is gratefully acknowledged.

A Quantitative Determination of the Neutron Moment in Absolute Nuclear Magnetons

LUIS W. ALVAREZ, *Radiation Laboratory, Department of Physics, University of California, Berkeley, California*

AND

F. BLOCH, *Department of Physics, Stanford University, Palo Alto, California*

(Received October 30, 1939)

The magnetic resonance method of determinng nuclear magnetic moments in molecular beams, recently described by Rabi and his collaborators, has been extended to allow the determination of the neutron moment. In place of deflection by inhomogeneous magnetic fields, magnetic scattering is used to produce and analyze the polarized beam of neutrons. Partial depolarization of the neutron beam is observed when the Larmor precessional frequency of the neutrons in a strong field is in resonance with a weak oscillating magnetic field normal to the strong field. A knowledge of the frequency and field when the resonance is observed, plus the assumption that the neutron spin is $\frac{1}{2}$, yields the moment directly. The theory of the experiment is developed in some detail, and a description of the apparatus is given. A new method of evaluating magnetic moments in all experiments using the resonance method is described. It is shown that the magnetic moment of any nucleus may be determined directly in absolute nuclear magnetons merely by a measurement of the *ratio* of two magnetic fields. These two fields are (*a*) that at which resonance occurs in a Rabi type experiment for a certain frequency, and (*b*) that at which protons are accelerated in a cyclotron operated on the nth harmonic of that frequency. The magnetic moment is then (for $J=\frac{1}{2}$), $\mu=H_b/nH_a$. n is an integer and H_b/H_a may be determined by null methods with arbitrary precision. The final result of a long series of experiments during which 200 million neutrons were counted is that the magnetic moment of the neutron, $\mu_n=1.93_5\pm0.02$ absolute nuclear magnetons. A brief discussion of the significance of this result is presented.

INTRODUCTION

THE study of hyperfine structure in atomic spectra has shown that a large number of atomic nuclei possess an angular momentum and a magnetic moment. Since, according to the theories of Heisenberg and Majorana, protons and neutrons are recognized as the elementary constituents of nuclear matter, their intrinsic properties and particularly their magnetic moments have become of considerable interest. The fundamental experiments of Stern and his collaborators[1] in which they determined the magnetic moments of the proton and the deuteron by deflections of molecular beams in inhomogeneous fields gave the first quantitative data of this sort. The approximate values which they gave for the two moments,[2]

$$\mu_p=2.5, \tag{1}$$

$$\mu_d=0.8, \tag{2}$$

suggested that in all likelihood, one would have to ascribe to the neutron a magnetic moment of the approximate value

$$\mu_n=-2. \tag{3}$$

The negative sign in formula (3) indicates that the relative orientation of their magnetic moments with respect to their angular momenta is opposite in the case of the neutron to that of the proton and the deuteron.

The technique of molecular beams has been greatly developed during the last few years by Rabi and his collaborators;[3] their ingenious methods have allowed them to determine the magnetic moments of many light nuclei with high precision, and to establish the existence of an electric quadrupole moment of the deuteron. Their values for the magnetic moments of the proton and deuteron are

$$\mu_p=2.785\pm0.02, \tag{4}$$

$$\mu_d=0.855\pm0.006. \tag{5}$$

They have also demonstrated that both moments

[1] R. Frisch and O. Stern, Zeits. f. Physik **85**, 4 (1933); I. Estermann and O. Stern, Zeits. f. Physik **85**, 17 (1933); O. Stern, Zeits. f. Physik **89**, 665 (1934); I. Estermann and O. Stern, Phys. Rev. **45**, 761 (1934).

[2] We shall throughout this paper give magnetic moments in units of the nuclear magneton $eh/4\pi Mc$ (e and M = charge and mass of the proton, c = velocity of light), and angular momenta in units $\hbar=h/2\pi$. The angular momentum of the deuteron is taken to be 1, that of the proton and the neutron to be $\frac{1}{2}$; the last is an assumption for which a direct experimental proof is still lacking, but which is generally accepted.

[3] I. I. Rabi, Phys. Rev. **51**, 652 (1937); I. I. Rabi, J. R. Zacharias, S. Millman and P. Kusch, Phys. Rev. **55**, 526 (1939).

are positive with respect to the direction of the angular momentum.

An experimental proof that a free neutron possesses a magnetic moment, and a measure of its strength, could also be achieved in principle by deflection of neutron beams in an inhomogeneous magnetic field. But while the great collimation required for this type of experiment may easily be obtained with molecular beams, it would be almost impossible with the neutron sources available at present. Better suited for the purpose is the method of magnetic scattering, which was suggested a few years ago by one of us.[4] It is based upon the principle that a noticeable part of the scattering of slow neutrons can be due to the interaction of their magnetic moments with that of the extranuclear electrons of the scattering atom. In the case of a magnetized scatterer this will cause a difference in the scattering cross section, dependent upon the orientation of the neutron moment with respect to the magnetization, and particularly in the case of ferromagnetics, it will cause a partial polarization of the transmitted neutron beam. The magnetic scattering of neutrons, and thereby the existence of the neutron moment, has been proved experimentally by several investigators, particularly by Dunning and his collaborators.[5] The magnetic scattering, however, can yield only a qualitative determination of the neutron moment since the interpretation of the effect is largely obscured by features involving the nature of the scattering substance.

Frisch, v. Halban and Koch[6] were the first to attempt to use the polarization of neutrons merely as a tool, and to determine the neutron moment by a change of the polarization, produced by a magnetic field between the polarizer and the analyzer. Such a change should indeed occur, because of the fact that the moment will precess in a magnetic field; by varying the field strength, one can reach a point where the time spent by the neutrons in the field is comparable to the Larmor period. In this way, one could obtain at least the order of magnitude of the moment. Although these investigators have reported an effect of the expected type, yielding the order of magnitude 2 for the neutron moment, we have serious doubts that their results are significant. Their polarizer and analyzer consisted of rings of Swedish iron, carrying only their remanent magnetism ($B=10{,}000$ gauss), while in agreement with Powers' results, we were never able to detect any noticeable polarization effects, independent of the kind of iron used, until it was magnetized between the poles of a strong electromagnet with an induction well above 20,000 gauss. Although we cannot deny the possibility that, due to unknown reasons, their iron was far more effective for polarization at low values of the induction than that used by other investigators, we think it more likely that in view of their rather large statistical errors the apparent effect was merely the result of fluctuations.

Although most valuable as a new method of approach, the experiment of Frisch, v. Halban and Koch could in any event give only qualitative results. The slow neutrons which one is forced to use emerge from paraffin with a complicated and none too well-known velocity distribution. The time during which they precess in the magnetic field will therefore be different for different neutrons and vary over a rather large range. Since it is that time which together with the field of precession determines the value of the moment, the latter will be known only approximately. A quantitative determination of the neutron moment therefore requires an arrangement which does not contain such features.

Method

Sometime ago, we conceived of an experimental method which could yield quantitative data of this sort. The method was independently proposed by Gorter and Rabi,[7] and most successfully used by the latter in his precision determinations

[4] F. Bloch, Phys. Rev. **50**, 259 (1936); **51**, 994 (1937); Ann. Inst. H. Poincaré **8**, 63 (1938); J. Schwinger, Phys. Rev. **51**, 544 (1937).

[5] P. N. Powers, H. Carroll and J. R. Dunning, Phys. Rev. **51**, 1112 (1937); P. N. Powers, H. Carroll, H. Beyer and J. R. Dunning, Phys. Rev. **52**, 38 (1937); P. N. Powers, Phys. Rev. **54**, 827 (1938).

[6] R. Frisch, H. v. Halban and J. Koch, Phys. Rev. **53**, 719 (1938).

[7] I. I. Rabi, J. R. Zacharias, S. Millman and P. Kusch, Phys. Rev. **53**, 318 (1938). During the summer of 1938 a first attempt to measure the neutron moment with such an arrangement was made at Stanford by Bloch, Bradbury and Tatel, using a 200-mg Ra-Be source of neutrons. However, it was unsuccessful due to instrumental effects which were hard to discover with the small intensity available.

of nuclear moments. Its principle consists in the variation of a magnetic field H_0 to the point where the Larmor precession of the neutrons is in resonance with the frequency of an oscillating magnetic field. The ratio of the resonance value of H_0 to the known frequency of the oscillating field gives immediately the value of the magnetic moment.[8]

The observation of the resonance point is based upon the fact that in its neighborhood there will be a finite probability P for a change of the orientation of the neutron moment with respect to the direction of the field H_0. Let this field be oriented in the z direction and let there be perpendicular to it, say in the x direction, an oscillating field with amplitude H_1 and circular frequency[9] ω, so that the total field in which a neutron is forced to move, is given by its components

$$H_x=H_1\cos(\omega t+\delta);\quad H_y=0;\quad H_z=H_0. \tag{6}$$

The solution of the Schroedinger equation for a neutron with angular momentum $\frac{1}{2}$ and magnetic moment μ gives the probability that a neutron, which at time $t=0$ in such a field had a z component $m=\frac{1}{2}$ of its angular momentum, will be found at the time $t=T$ with a value $m=-\frac{1}{2}$, in the form

$$P=\sin^2[(\mu H_1T/2\hbar)\times(1+\{2\Delta H/H_1\}^2)^{\frac{1}{2}}]/[1+(2\Delta H/H_1)^2], \tag{7}$$

where

$$\Delta H=H_0-H_0^*=H_0-\hbar\omega/2\mu \tag{8}$$

is the difference between the constant field H_0 and its value at resonance,

$$H_0^*=\hbar\omega/2\mu, \tag{9}$$

for which the Larmor frequency $2H_0\mu/\hbar$ is equal to the frequency ω of the oscillating field. Since the time T which the neutrons spend in the oscillating field will, for different neutrons, vary over a wide range, it will be a good approximation to substitute for the $\sin^2$ in the numerator of (7) its average value $\frac{1}{2}$. This means that, at resonance, complete depolarization of an originally polarized neutron beam will occur, and leads to the simplified formula

$$P=\tfrac{1}{2}[1+(2\Delta H/H_1)^2]^{-1}. \tag{7a}$$

Formulae (7) and (7a) have been derived by neglecting in the differential equation under consideration certain rapidly oscillating terms with respect to slowly varying ones; this neglect is justified as long as the condition

$$H_1/H_0\ll 1 \tag{10}$$

holds. In our final experiments, H_1/H_0 was less than 2 percent; so we felt justified in applying formulae (7) and (7a).*

The dependence of P upon ΔH, and particularly its maximum value for which the field H_0 has its resonance value H_0^* given by Eq. (9), is detected by a polarizer-analyzer arrangement, similar to that used by Frisch, v. Halban and Koch. We shall show that a beam of neutrons passing through two successive plates of magnetized material will change its transmitted intensity by an amount proportional to P. Consider two plates F_1 and F_2 both magnetized in the $+z$ direction and call f_i^+ and f_i^- $(i=1,2)$ the fraction of neutrons with $m=+\frac{1}{2}$ and $m=-\frac{1}{2}$, respectively, transmitted by F_i. f_i^+ will be different from f_i^- since the magnetic scattering will be preferential for one of the two values of m, thus giving rise to a partial polarization of the neutrons. If there does not occur any change of polarization between F_1 and F_2, the fraction of the neutrons transmitted by both will be given by

$$f=f_1^+f_2^++f_1^-f_2^-. \tag{11}$$

If, on the other hand, there is a probability P of a change of polarization between F_1 and F_2, the transmitted fraction will be

$$f'=(1-P)(f_1^+f_2^++f_1^-f_2^-)+P(f_1^+f_2^-+f_1^-f_2^+). \tag{12}$$

One will thus observe a relative change ΔI of the intensity I of the transmitted neutrons, given by

$$\Delta I/I=(f'-f)/f=-\alpha P, \tag{13}$$

[8] P. N. Powers, reference 5, has employed a modification of this method to determine the sign of the neutron moment. Instead of encountering an oscillating field between polarizer and analyzer, his neutrons were acted upon by a field which appeared to them, in consequence of their velocity, to be rotating with approximately their Larmor precession frequency. The sign of the moment was found to be negative, in agreement with theory.

[9] We shall always use circular frequencies although we sometimes refer to them merely as "frequencies."

* The corrections, introduced by this neglect are by no means of the order of 2 percent, but very much smaller, since they are essentially given by $(H_1/H_0)^2$. The mathematical theory of these corrections will soon be published.

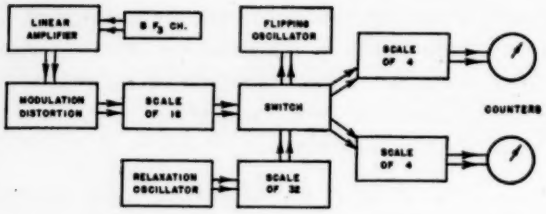

FIG. 1. Schematic drawing of electrical circuits.

with

$$\alpha=[(f_1^+-f_1^-)(f_2^+-f_2^-)]/(f_1^+f_2^++f_1^-f_2^-), \quad (14)$$

where α is a constant, characteristic of the magnetic and geometric conditions of F_1 and F_2 only. We see that this constant α can also be determined independently of any possible changes of polarization. For two identical plates F_1 and F_2 with common thickness α and with the values $f_1^{\pm}=f_2^{\pm}=e^{-(\sigma\pm\Delta\sigma)\alpha}$ ($\Delta\sigma=0$ for unmagnetized material) as they follow from the theory of magnetic scattering, one can see that α also has the significance of the relative increase of the total transmitted intensity which one observes by magnetizing F_1 and F_2 to the stage in which the fractions f_i have the values assumed in formula (14). Thus under ideal conditions, and if P assumes its maximum value $\frac{1}{2}$ at resonance, the relative change in intensity will just be $\frac{1}{2}$ of that observed when demagnetizing F_1 and F_2 in an experiment similar to that of Dunning and his collaborators.

Once the maximum value of (13) has been experimentally established, i.e., the value of H_0 has been found for which $\Delta H=0$, or where according to (8) and (9)

$$H_0=H_0^*=\hbar\omega/2\mu, \quad (15)$$

one will know the magnetic moment μ in the form

$$\mu=\hbar\omega/2H_0^*, \quad (15a)$$

provided that the values of H_0^* and the frequency ω of the oscillating field are known.

One can of course measure the latter by means of an ordinary standardized wavemeter and the former by the usual method of flipping a coil in the magnetic field and comparing the ballistic throw of a galvanometer with that produced by breaking the current in the primary of a known mutual inductance. We have, however, thought of an alternative method, which not only can be carried to almost unlimited accuracy, but also has the highly desirable feature of yielding the moment to be determined directly in units of the nuclear magneton. It consists in comparing the frequency $\omega=\omega_n$ and the field $H_0^*=H_n$ with the values ω_p and H_p, where ω_p is the circular frequency of an oscillator driving a cyclotron which accelerates protons in a magnetic field H_p. We have them from (15)

$$\omega_n=(2H_n\mu_n/\hbar)(e\hbar/2Mc)=(eH_n/Mc)\mu_n, \quad (16)$$

if instead of using absolute units as in (15), we express the neutron moment in units of the nuclear magneton, i.e., if we write

$$\mu=\mu_n(e\hbar/2Mc). \quad (17)$$

On the other hand, for the ideal resonance of protons in a cyclotron, we have the relation

$$\omega_p=eH_p/Mc. \quad (18)$$

Dividing (16) by (18), we now obtain

$$\mu_n=(\omega_n/\omega_p)(H_p/H_n). \quad (19)$$

This means that in order to determine the moment in absolute nuclear magnetons one has merely to determine the *ratio* of two frequencies and the *ratio* of two fields instead of determining ω_n and H_n absolutely. In addition to this great experimental advantage, we see that no knowledge of the universal constants of physics is required. The value of Planck's constant does not enter since the moment is expressed in nuclear magnetons and the value of e/Mc is eliminated by use of the cyclotron.

If the radiofrequency generator for the cyclotron were a power amplifier driven by a harmonic of the "neutron oscillator," the ratio ω_n/ω_p would be simply an integer. In this case one would merely have to measure the ratio H_p/H_n, in order to know μ_n. This measurement can be reduced to the determination of the ratio of the wound areas (area $\times$ turns) of two flip coils, which can in turn be reduced to that of two currents. If the two coils, located in the same position in a current system and both connected through the same galvanometer circuit give the same ballistic throw, then the ratio of the currents for which this condition is reached gives the ratio of their wound areas. Once that ratio is known, one also has that of any two fields which,

using the two coils, give the same ballistic throw in the galvanometer. Gehrcke and v. Wogau[10] have developed this technique, by the use of null indicator methods, to a point where the ratio of two fields may be determined to five significant figures. Although we have determined the ratios of the frequencies and the fields by essentially these methods the errors, entering from elsewhere did not justify the application of all their refinements described above.

Description of Apparatus

Up to the present time, almost all of the precision measurements of neutron intensity have been made using Rn-Be, or similar sources not subject to the erratic fluctuations occurring with artificial sources. Although it is possible to keep the cyclotron ion beam constant to one or two percent for long periods of time, this is not sufficient for the present work, where an accuracy of the order of 0.1 percent is needed. Also, one cannot assume that the neutron intensity is proportional to the beam current, since the distribution of the latter over the target, and the proportion of stray ions striking the insides of the dees are both important in this respect. This rules out the possibility of using an integrating beam meter, or auxiliary BF_3 chamber, as a monitor. The counting-monitoring circuit which was finally adopted for this experiment is shown schematically in Fig. 1.

The neutrons were detected in a BF_3 chamber operated at 5000 volts. Its diameter and length were 9 cm and 15 cm, respectively, and the gas pressure was 50 cm of Hg. The arrangement of electrodes is quite similar to that of a chamber recently described by Powers.[3] A conventional linear amplifier of 4 stages fed the pulses into the modulation-distortion circuit, where all pulses of magnitude greater than a certain value were amplified to uniform size, and all below this height were suppressed. This circuit has been described in detail before, in connection with the production of monochromatic neutron beams,[11] but was not essential for most of this work, since modulated beams were not used in the final experiment. The uniform pulses from this stage were fed to a scale-of-16 of the type described by Reich,[12] which uses hard tubes. After passing through the scale, the pulses entered a switching circuit which fed them to either of two scales of 4 plus a Cenco counter. The switching was carried out automatically at regular intervals of from 2–10 seconds, depending upon the adjustment. At the same time, the switching circuit turned a magnet on and off (from magnetic scattering experiments), or keyed the oscillator (for the resonance experiment).

The first circuits designed to carry out the switching used relays; these were found to be unsatisfactory due to differences in spring adjustments, so an all electronic switching circuit was devised. This is shown in Fig. 2. Negative pulses were fed to the grids of the two right-hand 6K7's, each of which had another 6K7 connected to it in the familiar Rossi coincidence circuit. If the grids of either of these left-hand 6K7's were negative, the plate current of its partner would be cut off, delivering a positive pulse to the corresponding scale-of-4. But if the grid was positive, the voltage pulse at the common plate connection would be too small to affect the scale. The circuit at the left is merely a device for making one grid positive and the other negative, and reversing this condition instantaneously at will. The two 885's are arranged in a conventional inverter, or scale-of-2 circuit. When current flows through the upper Thyratron and is blocked in the lower, pulses will pass through the switch into the upper scale-of-4. If a positive pulse is now applied to the grids of the 885's, the scale-of-2 will invert, and the lower scale-of-4 will receive the neutron pulses. The positive pulse which flips the scale-of-2 originates in a relaxation oscillator of the type used in cathode-ray oscillograph sweep circuits. It was found that the interval between

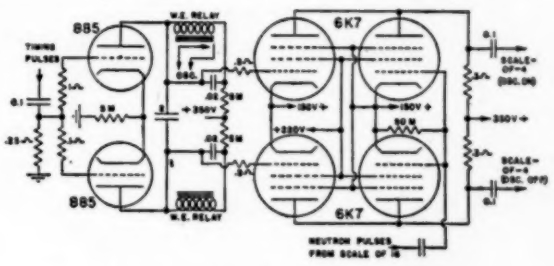

Fig. 2. Electronic switching circuit.

[10] E. Gehrcke and M. v. Wogau, Verh. d. D. Phys. Ges. **11**, 664 (1909).
[11] L. W. Alvarez, Phys. Rev. **54**, 609 (1938).
[12] H. Reich, Rev. Sci. Inst. **9**, 222 (1938).

pulses was not constant when R and C were large enough to make the time constant 5 seconds; so the frequency was increased, and the pulses from the oscillator were fed through a scale of 32 into the switching circuit. Fluctuations in the time interval were reduced to a negligible amount by this procedure. The length of the interval could be varied by means of a variable resistance in the oscillator circuit. Small relays in the plate circuits of the Thyratrons made it possible to control the magnets or R.F. oscillator in phase with the counter switching. A lag in these operations merely reduced slightly any real effect, while a lag in the counter switching would introduce a spurious effect of the same magnitude.

On first thought it would seem that the statistical errors would be considerably increased by the method of switching a scale-of-16 from one scale-of-4 to another. But calculation of the effect shows that it is of no importance, and our experimental fluctuations have always checked with those to be expected on the assumption that this arrangement is equivalent to two separate scales-of-64. We have also made runs to test the equality of the counting time under the two conditions, by feeding the output of a beat frequency oscillator into the distortion amplifier. It was always found that the two Cenco counters recorded equal numbers of counts per cycle of the switching.

We spent several months developing our polarizer-analyzer equipment. Our first attempts were along lines similar to those used by Frisch, v. Halban and Koch. Flat rings of iron with a diameter of 1 meter, a thickness of 1 cm, and a width of 10 cm were wound with copper wire in toroidal fashion to minimize the stray field. We were unable to observe any difference in the transmitted neutron intensity, which depended upon the magnetization of the iron. This is in agreement with the Columbia school which has shown that this iron (Armco) has a sharp magnetic effect threshold very near the saturation value, which most likely was not reached with our low magnetizing currents.

Since they also report that Swedish iron shows neutron magnetic effects at lower fields, we next wound bars of Swedish iron with flat copper strip through which 120 amperes could be passed. The bars were 1.5 meters long, 10 cm wide, and 1.5 cm thick. Although the H in the iron was 250 gauss ($B=20{,}000$), we were still unable to observe a difference in neutron transmission. We finally cut a piece from one of the Swedish iron bars, and magnetized it strongly between the pole pieces of a Weiss magnet and observed the magnetic effect. With an iron piece 4 cm thick, we obtained differences of about 6 percent in transmission. With half of the iron in the Weiss magnet, and the other half in a similar strong magnet, we confirmed Powers' results on double transmission, i.e., that the change in intensity was about the same for the pieces in one magnet, as when they were separated in two, at some distance. Many experiments were made to determine

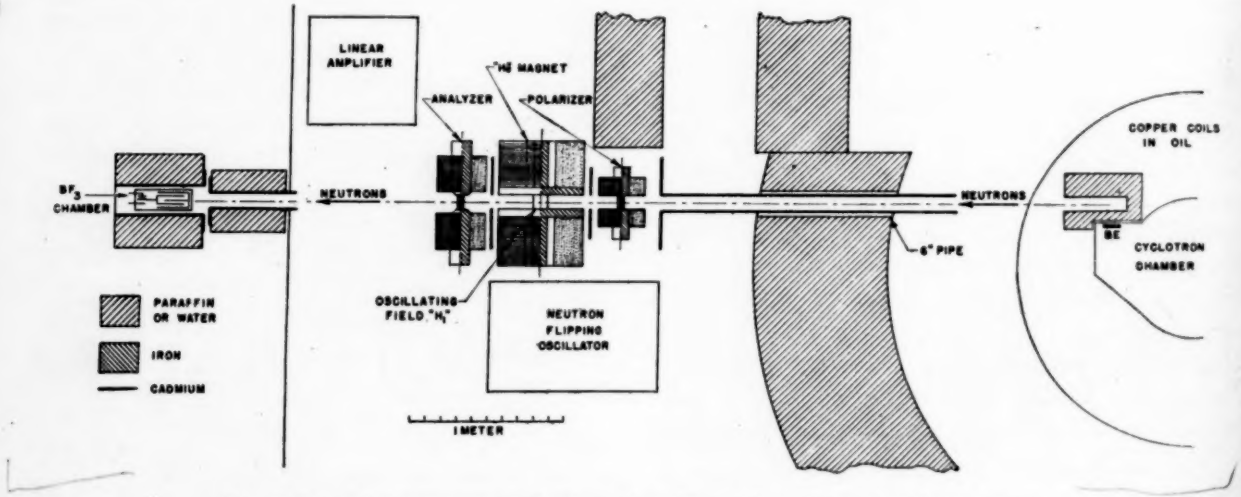

FIG. 3. Plan of the apparatus. Section taken 85 cm above floor of cyclotron room.

the optimum thickness of iron, and the magnetizing currents needed to obtain the maximum effects.

It was then possible to consider the construction of the precessing field, and the oscillator system. A plan of the apparatus is given in Fig. 3. Neutrons from the Be+D reaction were slowed to thermal velocities in the simple howitzer, and diffused down a cadmium-lined tube through the water tank to the polarizer. After passing through the precessing and flipping field, and the analyzer magnet, they were recorded in the BF_3 chamber. The large blocks of paraffin which defined the beam made the fast neutron background quite small. The iron plates reduced the thermal neutron intensity to about 5 percent, but the fast neutron background, as measured by covering one of the plates with Cd, was still only 25 percent of the total counting rate. This small background made it unnecessary to use the backgroundless modulated beam, as was originally planned. The somewhat increased magnetic effect at low neutron velocities was not enough to offset the decrease in intensity which attended the use of the modulation. In the earlier stages of the experiment, and actually in one of the resonance experiments, 120-cycle modulation was used successfully, but it was abandoned in favor of the greater intensity available without modulation.

The oscillator was of the conventional Hartley type, using an Eimac 150*T*, self-excited. The frequency was found to remain constant to better than 0.05 percent for days. The solenoid which produced the oscillating magnetic field was wound with a 9-cm square cross section and 15 cm height. The conductor was flat copper strip $0.50''\times0.010''$ wound so that adjacent turns almost touched. Neutrons passed through the strip on entering and leaving the region of the oscillating field H_1. The solenoid was connected in parallel with a section of the tank coil chosen to give 11 amperes magnetizing current as measured on an R.F. ammeter in the solenoid circuit ($H_1\approx10$ gauss). The oscillator was keyed by means of a relay in the primaries of the plate supply transformers.

The steady precessing field H_0 was at first produced by a pair of Helmholtz coils plus a pair of cylindrical bars of iron to concentrate the flux. This field was quite inhomogeneous, and was used for the first two resonance dips only. For the later measurements, we used a pair of large coils which had energized the magnet of the original 1-mv cyclotron of Lawrence and Livingston. These magnet coils were loaned to us by Professor R. B. Brode, who had been using them in connection with a cosmic-ray cloud chamber. An arrangement of iron was found which produced a field of about 600 gauss ±1 gauss throughout the volume occupied by the R.F. solenoid. The fluctuations in the laboratory d.c. supply were too great to allow its use in the coil circuit, so we utilized the constant current supply of the cyclotron magnet. This automatic stabilizer keeps the voltage fluctuations across the cyclotron magnet coils to less than 0.1 percent and gradually increases the value of the stabilized voltage to compensate for the heating of the coils. The magnet coils were then connected in parallel with the cyclotron coils through a variable resistance. The current was watched constantly, and the resistance was varied slightly when necessary, to allow for the differential heating of the two sets of coils.

A pair of large grounded copper plates were placed against the polepieces, so that the magnet coils were shielded from the R.F. fields of the solenoid. If this shield was not grounded, radio-frequency currents were induced in the d.c. lines and fed back into the stabilizer. A partial rectification of these currents would cause the stabilizer to change the magnet current, and therefore throw the cyclotron out of resonance. The resulting drop in the neutron intensity would be in phase with the switching cycle, and would therefore produce an apparent effect of the same sign and magnitude of that to be expected. Another cause of apparent effects was the interaction of the R.F. oscillator with the relaxation oscillator in the timing circuit. When the scale-of-32 was not properly grounded, it would count an extra pulse whenever the oscillator was turned on. This gave a 3 percent apparent effect of the correct sign and magnitude, due to the change in switching period. Several other spurious effects were encountered at various times during the experiment; it is due to the large intensities of neutrons available that we were able to trace

down and eliminate them after comparatively short counting times.

The test for reliability of the apparatus in which we placed most faith was the following: A series of observations was taken at 20μa of deuterons, and it was desired to know whether the effect, e.g., difference in neutron counts with and without the oscillator on, was real. The current was then increased to 80 amperes, and the analyzer was covered with a cadmium sheet. The counting rate in all parts of the amplifier, from the BF_3 chamber to the final mechanical counters remained the same as before, and all relays, magnetic and electric fields were operating in the same manner. The only difference was that the counted neutrons did not for the most part pass through the polarizer and analyzer, and that part of high velocity which did could not exhibit any magnetic effects. Any difference between the relative counting rates under these two conditions was ascribed to magnetic effects of the thermal neutrons. It was also possible in this way to eliminate causes of systematic effects. The closing of relays would occasionally jar the BF_3 chamber, and the resulting noise would introduce extra counts in synchronism with the timing cycle. Several effects of this nature were discovered and eliminated after long blank runs. It was also important at these high counting rates to guard against saturation in the BF_3 chamber. The intensity was always adjusted so that the counting rate was still a linear function of the deuteron current.

These sources of trouble are mentioned to emphasize the great desirability of powerful neutron sources in experiments of this sort. In the final set of experiments, in which all these spurious effects were eliminated and on which our value of the neutron moment is based, only about 4 million neutrons were counted. This number of counts could be taken with Rn-Be sources of moderate strength in a reasonable length of time, but it would be almost impossible to track down obscure systematic effects in the manner outlined above. We counted more than 200 million neutrons in test runs before we felt sure enough of the apparatus to have confidence in the results outlined below. Perhaps the most convincing proof of the reality of the effect was that we varied H_0 and ω over a factor of almost 10, and *each time the frequency was increased, the magnetic field setting for the center of the new dip was found within the accuracy, to be expected at the value calculated from the previously determined value of* H_0^*/ω.

Measurements

In the first experiments with the small Helmholtz coils, the magnetic field was measured with the aid of a fluxmeter-search coil combination. Although somewhat crude, the measurements did show that H_0^*/ω was constant for two frequencies differing by about 20 percent, and that the effect could therefore be attributed to the magnetic moment of the neutron, as outlined above. The shallowness of the resonance dip observed in these first experiments ($\Delta I/I=0.6$ percent) could be attributed to the inhomogeneity of the field H_0.

When the large magnet coils were installed, it was clear that a more accurate method of measuring the field would be necessary. A brass flip coil was therefore constructed in such a way that it could be alternately lowered into a standard position within the R.F. solenoid, or withdrawn to allow the neutrons to pass through. Stops were placed so that the coil turned through exactly 180°, as determined by optical means. The coil was turned by spring tension, after being cocked by hand. The leads from the coil were connected in series with a ballistic galvanometer of very long period, a Hibbert magnetic standard, and a search coil at the center of a standard solenoid. A knowledge of the constant of the standard solenoid plus the wound area of the flip coil and the search coil allowed the strength of the field to be ascertained. The current through the standard solenoid was measured by means of a standard resistance, standard cell, and potentiometer.

Before and after each counting period, the relation between the current through the coils, and the magnetic field was determined. This consisted in noting the ratio of the throws of the ballistic galvanometer d/s, where d was the deflection due to the turning of the flip coil, and s was the throw of the magnetic standard. Effects due to any nonlinearity of the galvanometer were minimized by adjusting this ratio close to one by rewinding the standard. d/s was determined as a

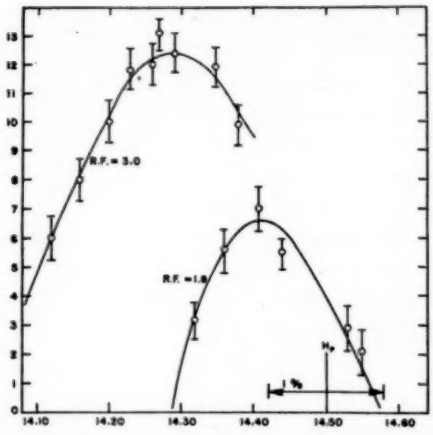

FIG. 4. Proton resonance curves. The potential drop in millivolts across the magnet current shunt is plotted as abscissa against the electromotive force in millivolts of the thermocouple used in the probe on which the proton current fell.

function of the magnetizing current, and the neutron intensity was also measured as a function of this current. The resonant field was then found from a comparison of these two curves.

The frequency of the oscillator was measured on a General Radio precision wavemeter. A knowledge of the field and frequency yielded the magnetic moment of the neutron by substitution in Eq. (15a).

The most trustworthy value of the moment is, however, given in absolute nuclear magnetons by the comparison method, outlined above. Although we ultimately trusted more this method it was gratifying to obtain well within the experimental error the same value for μ_n by the two entirely different methods of calibration. For this purpose it was necessary to construct a new flip coil which when turned through 180° in the proton resonance field of the 60-inch cyclotron would give a d/s almost equal to that obtained in the measurement of H_0*. From a knowledge of the approximate values of the two magnetic fields involved, and the wound area of the original coil, it was a simple matter to construct such a coil. This coil was wound on a hard rubber form to ascertain the absence of magnetic impurities in it or in the brass used for the first coil.

It was then necessary to determine the ratio of the wound areas of the two flip coils. This measurement was carried out in the following manner: The two coils were mounted near the center of a pair of carefully constructed Helmholtz coils, which were designed for accurate cloud-chamber beta-ray curvature measurements. The constant of the coils did not enter the calculation; it was merely assumed that the field was directly proportional to the current through the coils. This current was measured in terms of the potential drop across a standard resistance in the circuit, on a Wolff potentiometer standardized with a Weston cell. The small coil was flipped in the strongest field available, and the current was then reduced until the large coil gave the same ballistic throw as the small one in the higher field. (The effect of the earth's field was subtracted from all readings.) The ratio of the wound areas of the two coils was then given by the ratio of the currents as determined on the potentiometer. The two coils were connected in series with the galvanometer; so the total resistance of the circuit remained constant. The current ratio was found to be constant for many values of the fields. To test for magnetic impurities in the search coil, the coil was connected in series opposition, and the large coil was turned through such an angle that when a current was broken in the Helmholtz coils no kick was observed on the galvanometer. This condition prevailed for a wide range of magnetizing currents. A very sensitive galvanometer was used in these measurements; so this zero method gave us confidence that no appreciable ferromagnetic contamination existed in the flip coils, one of brass and the other of hard rubber.

The new coil, ballistic galvanometer, and magnetic standard were then carried to the Crocker Radiation Laboratory, where the magnetic calibration was made in terms of the 60-inch cyclotron field. The rod carrying the flip coil projected into the cyclotron chamber through a window so that the coil could rotate in the gap at a point about 25 cm from the center of the cyclotron. Tests with a large search coil connected to a fluxmeter showed the field to be constant throughout this central region to better than 0.1 percent. The magnetization curve of the cyclotron magnet was then measured in the region around the known proton field. The measurements consisted in a determination of d/s against the potential drop across the magnet current shunt, as measured on the Wolff potentiometer.

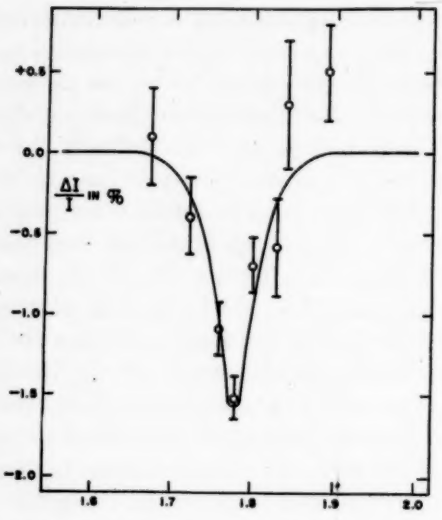

FIG. 5. Neutron resonance dip. The magnet current in arbitrary used is plotted against the fractional change ($\Delta I/I$) of the intensity of the neutron beam.

The cyclotron was then evacuated and hydrogen let into the chamber. The heating effect of the proton current to a probe placed at the point occupied by the flip coil was then measured as a function of the potential drop across the magnet current shunt. A comparison of these two curves gave the d/s value for the acceleration of protons at the frequency in use at the time. The effects of hysteresis on the resonance curve were found to be negligible, by observing it when the magnet current had been increased from zero to the proton value, and also when it had been carried through the cycle—zero, deuteron-field, proton-field.

The proton resonance curve, Fig. 4, was measured at two values of the dee voltage, and the center of gravity of the curve was found to move toward higher fields as the voltage was reduced. This was shown not to be caused by a change in the cyclotron frequency, by measurements taken during each run. The increase in width of the curves with voltage is due to the fact that at higher voltage, the protons have to circulate fewer times. The shift of the peaks is due mainly to the focusing properties of the cyclotron. Since the magnetic field was uniform out to the probe, the focusing was entirely electrostatic, and this is known to be dependent upon the phase of the protons with respect to that of the oscillator.[13]

[13] M. E. Rose, Phys. Rev. **53**, 392 (1938); R. R. Wilson, Phys. Rev. **53**, 408 (1938).

Professor E. McMillan and Mr. R. Wilson have made a study of these effects in this laboratory, and they agree that the proper method of choosing the magnetic field, H_p, would be to run a large series of resonance curves with decreasing dee voltage, and extrapolate the resonance field to zero dee voltage. Changes in the apparatus made this impossible, so they have suggested that we choose a point midway between the maximum of the curve taken at lower dee voltage and its high field intercept. Since we have seen both experimentally and theoretically that the extrapolated resonance curve will lie somewhere between these two points, separated in field by 1 percent, the maximum error introduced in this way could be only 0.5 percent.

The ratio of the cyclotron frequency to the neutron frequency was determined with the aid of a Super-Pro short wave radio receiver. The second harmonic of the cyclotron oscillator was tuned in on this receiver in the same frequency range as the sixth to tenth harmonics of the neutron oscillator. The agreement between the various values of the frequency of the neutron oscillator as determined from these measurements attests to the linearity of the frequency scale of the receiver. The receiver was thus used merely as an uncalibrated, linear wavemeter.

RESULTS

Since the width of the resonance dip in gauss is constant for all values of the precessing field (Eq. (7a)), it is obvious that it is desirable to

TABLE I. *Data for neutron resonance curve (Fig. 5). Each reading is based on 120,000 counts. Magnetic field for neutron resonance, 622 gauss; frequency of oscillator, 1843 kilocycles; magnetic moment of the neutron, 1.94 nuclear magnetons; d/s for coil 1 in neutron field, 0.866; d/s for coil 2 in proton field, 0.844.*

CURRENT	PERCENTAGE EFFECT	MEAN EFFECT
1.67	−0.8, +1.0	+0.1 ±0.3
1.72	+0.6, −0.9, −0.8	−0.4 ±0.23
1.76	−0.5, −0.9, −2.1, −1.1, −1.2, −0.8	−1.1 ±0.16
1.78	−2.0, −2.2, −1.6, −1.5, −1.4, −0.5, −1.7, −1.2, −1.6, −1.5	−1.52±0.13
1.79	−1.7, −1.7	−1.7 ±0.3
1.80	−1.4, −0.5, −0.7, −1.5, +0.1, −0.3	−0.71±0.16
1.81	−1.5	−1.5 ±0.4
1.83	−0.8, −0.4	−0.6 ±0.3
1.84	+0.3	+0.3 ±0.4
1.89	+1.0, +0.0	+0.5 ±0.3

perform exploratory experiments at low precessing fields, to find an approximate value of the magnetic moment. After the relatively wide dip has been found at low fields, it is possible to increase the precision of the determination by raising the field and the frequency in the same ratio. Our first measurements were made at 260 and 315 kc, and merely served to show that the value of the moment was close to 2 nuclear magnetons. When the large homogeneous field was installed, we repeated the measurement at 315 kc; after observing the dip again at 810 kc, we increased the frequency to 1843 kc, where the field calibrations were made. Since the resonance dips at all the lower frequencies were consistent with that at 1843 kc, we will not describe those measurements in detail. The curve at 1843 kc is shown in Fig. 5.

The experimental data from which this curve is plotted are given in Tables I, II and III. It will be noticed that two of the points were not plotted in the curve. These are the values at 1.79 and 1.81 amperes, which have very large probable errors, due to the small numbers of neutrons counted. These two points have no appreciable effect upon the position of the minimum, and although they lie far off the curve, their positions are quite consistent with their probable errors. The curve was obtained on two separate days a week apart, and since no systematic difference could be noted between the two sets of data, they have been plotted together. The value of the magnetic moment is then given from Eq. (19),

$$\mu_n=(1.856/9.685)\times(10.60\times 0.844/0.886)$$
$$=1.935 \text{ nuclear magnetons.}$$

The *probable errors* entering from various sources are: position of dip, 0.5 percent; galvanometer readings, 0.2 percent; coil calibration, 0.2 percent; frequency calibration, 0.2 percent; and cyclotron calibration, 0.6 percent. The final result may then be expressed as

$$\mu_n=1.93_5\pm 0.02$$

absolute nuclear magnetons.

TABLE II. *Current ratios in cloud-chamber Helmholtz coils for equal galvanometer throws.*

CURRENT	RATIO	CURRENT	RATIO
2.20	10.61	1.60	10.63
2.19	10.61	1.30	10.58
2.18	10.58	Ave.	10.60
2.10	10.60		

TABLE III. *Frequency calibration: reading, R, on radio receiver for nth harmonic.*

NEUTRON OSCILLATOR			CYCLOTRON OSCILLATOR		
n	R	R/n	n	R	R/n
6	11.14	1.856	2	19.37	9.685
7	12.99	1.856			
8	14.83	1.854	All values in megacycles.		
9	16.70	1.856			
10	18.56	1.856			

DISCUSSION

The now rather accurately known values

$$\mu_p=2.78_5\pm 0.02 \quad \mu_n=-1.93_5\pm 0.02$$
$$\mu_d=0.855\pm 0.006$$

of the magnetic moments of proton, neutron and deuteron are of considerable interest for nuclear theory. The fact alone that μ_p differs from unity and μ_n differs from zero indicates that, unlike the electron, these particles are not sufficiently described by the relativistic wave equation of *Dirac* and that other causes underly their magnetic properties.

Whatever these causes may turn out to be one has to notice that there holds to well within the experimental error the simple empirical relation

$$\mu_d=\mu_p+\mu_n. \tag{20}$$

This relation is far from being obvious and it would in fact seem rather surprising if it were rigorously satisfied. To explain it in simple terms one would have to make both the following assumptions:

(*a*) The fundamental state of the deuteron is a ${}^3S$ state so that there are no contributions to μ_d arising from orbital motion of the particles.

(*b*) The moments μ_p and μ_n are "additive," i.e., their intrinsic values are not changed by the interaction of the proton and the neutron, forming a deuteron.

The first assumption has been disproved by the recent discovery[14] that the deuteron possesses a finite electric quadrupole moment which is

[14] J. M. B. Kellogg, I. I. Rabi, N. F. Ramsey and J. R. Zacharias, Phys. Rev. **55**, 318 (1939).

incompatible with the symmetry character of a pure ^3S state. The second cannot be discarded on an experimental basis but it ceases to be plausible if one admits the possibility, that ultimately the same causes may underly both the magnetic properties and the mutual binding forces of the proton and the neutron.

It is conceivable that the departure from any one of the two assumptions (*a*) and (*b*) would separately cause a considerable deviation from (20) but that for unknown reasons both together cancel each other very closely. Until reliable estimates of these deviations can be obtained we consider it, however, more likely that neither of them amounts to more than a few percents.

ACKNOWLEDGMENT

We are indebted to Professor E. O. Lawrence for his interest in this experiment, and in particular, one of us is grateful to him for the opportunity to work as a guest in the Radiation Laboratory. Our sincere thanks are due Professor R. B. Brode for the loan of the magnet which made the precision determination of the magnetic moment possible. We also wish to express our appreciation to the members of the laboratory staff for their many hours at the cyclotron controls. The experiment was made possible by grants to the laboratory from the Rockefeller Foundation and the Research Corporation.

JANUARY 15, 1940 PHYSICAL REVIEW VOLUME 57

The Contact Difference of Potential Between Barium and Zinc

The External Work Function of Zinc

PAUL A. ANDERSON
State College of Washington, Pullman, Washington

(Received November 13, 1939)

The methods of contact potential measurement described in previous reports of this series have been extended to zinc, for which the published work function data are discordant. Measurement was by the retarding potential method with variable anode, in tubes sealed from the pumps and gettered with vaporized barium. Barium was the reference metal and the measured surfaces of both barium and zinc were prepared by fractional distillation followed by revaporization of the middle fractions. In observations on three tubes and fifteen pairs of Ba-Zn surfaces the observed contact P.D., mean value 1.76 v, was reproducible and constant to ±0.02 v. Combination of this value with the work function of barium, now well established at 2.52 ev, assigns the value 4.28±0.02 ev to the work function of zinc at room temperature. The initial measurement on a surface was taken within 8 seconds of its deposition. As a consequence of this small time lapse it becomes possible to answer the fundamental but neglected question of whether a work function measured under optimum vacuum conditions is characteristic of the clean metal or of a gas-contaminated surface in equilibrium with the residual gas. The rate of deposition of oxygen on zinc is estimated from the behavior of a tungsten "test surface" in a comparable vacuum. Since the regimes of measurement and of equilibrium gas film formation overlap and no change of work function with aging is detectable, it is concluded that the observed work function is characteristic of clean zinc. From the point of view of the adsorption processes involved, the measurements show that the mean time of sojourn of oxygen is much less on zinc than on tungsten.

THIS paper reports further results in a program of study of the external work functions of the pure metals as determined by measurement of their contact differences of potential with respect to a suitable reference metal, barium. Previous work under this program[1] has shown, (1) that the work functions of barium surfaces are reproducible to at least ±0.01 ev when the surfaces are formed by fractional distillation in a sealed-off tube, and (2) that measurements of work functions by the contact potential method are now capable of the precision and accuracy characteristic of the best photoelectric work. Since the contact potential method, unlike the thermionic, is applicable to all the metals and, unlike the photoelectric, yields measurements of uniformly high sensitivity, it appears to offer the most promising

[1] P. A. Anderson, Phys. Rev. **47**, 958 (1935); **49**, 320 (1936); **54**, 753 (1938).

road to a systematic knowledge of the work-function characteristics of noncomposite metal surfaces.

For the purpose of testing our method of contact potential measurement, previous studies have been made upon metals for which the photoelectric data are especially reliable. The method is now being extended to metals for which the photoelectric data are contradictory, or for which no work function studies have yet been made under adequate experimental conditions. Zinc is one of the numerous common metals for which these data are discordant. Published values[2] range from 3.08 to 4.24 ev and the agreement has scarcely been improved by the most recent work. DeVoe,[3] using the photoelectric method, has obtained 4.24 ev for the Fowler threshold while Oatley,[4] employing a contact potential method, has reported a work function of 3.40 ev. Differences in the crystal structure of the zinc surfaces, determined by variations in the methods employed by different investigators for preparing their specimens, are no doubt partly responsible for the divergent values reported, but structural differences cannot account for work function deviations greater than a few decivolts. As has been pointed out in previous papers of this series, it is desirable that a standardized technique be selected for the preparation of each metal which is subjected to study, that the work function of this standard surface be determined under the most rigorous vacuum conditions, and that studies of the structural dependence of the work function be then referred to this standard surface. It has been suggested, also, that in work on most of the common metals, standard surfaces of reproducible structure can be best prepared by fractional distillation of the metal in a gettered vacuum, with deposition on a given substrate at given substrate temperature. This choice of technique, borne out by several extended series of measurements in this laboratory, has received further support from photoelectric measurements made at Northwestern by Cashman and his associates, who have developed a similar technique. Recent results on magnesium films prepared by this standardized method and measured in independent photoelectric and contact-potential studies have given work function values which agree to 0.01 ev.[5] It was the purpose of the present work to extend this method to the

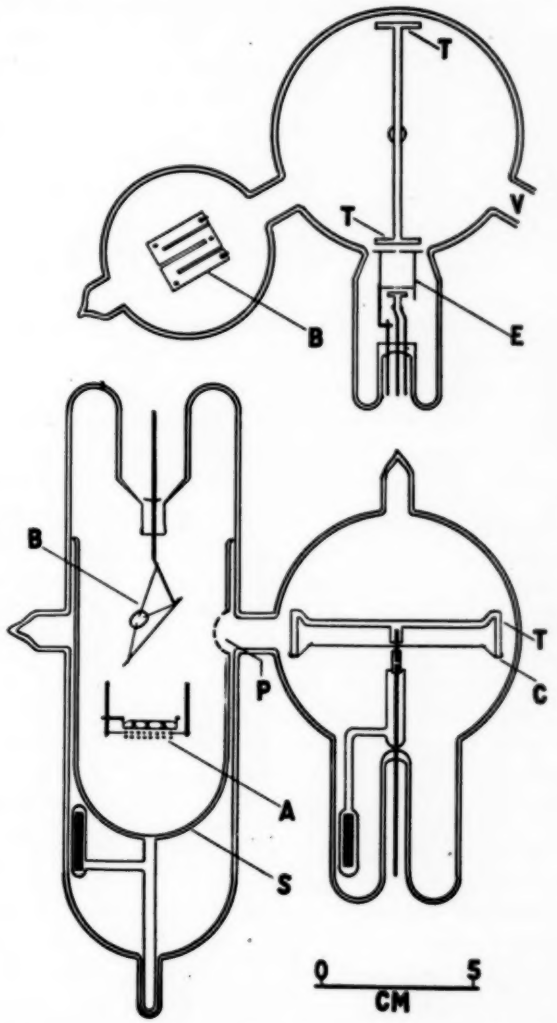

FIG. 1. The tube. First and second vaporizers, *A* and *B*, are mounted on single 3-lead press, one of 60-mil leads common. *A* (rotated 90° in sketch) comprises 20×20×5-mm pan of 1-mil Mo foil, with flat helical heater of 20-mil W wire. *B* is single 25×30-mm sheet of foil; Mo foil for barium, W foil for Zn (see text). Target assembly *TT* is of glass; rotatable on 60-mil W lead by magnetic traction on glass-enclosed armature. Contact *C* is 10-mil Pt wire embedded in face of target. Vapor shield *S*, with porthole *P*, is of glass with glass-enclosed armature. Electron gun *E* of 4-mil Ta foil; emitter 1-mil W foil 1.5 mm wide. A second vaporizing bulb, connected at *V*, is identical in design to the bulb shown.

[2] For a bibliography of the earlier work see A. L. Hughes and L. A. DuBridge, *Photoelectric Phenomena* (McGraw-Hill, 1932), p. 76.
[3] C. F. DeVoe, Phys. Rev. **50**, 481 (1936).
[4] C. W. Oatley, Proc. Roy. Soc. **A155**, 218 (1936).
[5] R. J. Cashman, Phys. Rev. **54**, 971 (1938); P. A. Anderson, Phys. Rev. **54**, 753 (1938).

determination of the contact difference of potential between barium and zinc. Both metals were fractionally and multiply distilled and condensed on glass at room temperature.

Experimental Procedure. Tube Design

The general features of the method and the electrical circuits have been described in previous reports. Details peculiar to the present measurements may be described in connection with the design of the tube, Fig. 1. The main improvement over our previous tubes lies in the construction of the first and second vaporizers and in their thorough shielding to eliminate any possibility of contamination of either metal by the other during the vaporizations. The first vaporizer *A*, providing for rapid fusion and controlled vaporization of a large charge (5–10 g) of metal by radiation heating, was an adaptation of a well-known method. The second vaporizer *B* is of new design which provides (1) a condensing surface of large area with consequent high efficiency in terms of the amount of metal finally utilized to that available in the middle fractions, and (2) very exact control of the rate at which vapor is delivered by the second vaporizer to the final condensing surface. This vaporizer was made by converting a 25×30-mm sheet of 1-mil tungsten or molybdenum foil into a grid by parallel cuts with a razor blade, as indicated at *B* in the plan sketch of Fig. 1. The grid was connected to its 60-mil tantalum current leads with tantalum rivets and held taut by a simple spider of 10-mil tungsten wires embedded in a Nonex bead and hooked through small holes in the grid, as indicated in the elevation sketch of Fig. 1. Heating current for the grid was supplied by a step-down transformer with continuously variable secondary voltage (Variac 80-A). The tubular shield *S*, mounted on a simple glass-to-glass thrust bearing as indicated, could be spun by a small permanent magnet and its porthole *P* brought into or out of register with the tube connecting vaporizing and measuring bulbs. During the preliminary fusions and vaporizations of barium and zinc both shields were kept closed; a shield was opened only for the deposition of a film on the target *T* and both shields were never opened simultaneously. In addition to their primary functions, the shields furnished an invaluable means of observing the progress of the vaporizations after the inner walls of the vaporizing bulbs had become coated with opaque layers of distilled metal. Slight rotation of a shield brought its porthole opposite an uncoated section of the outer bulb. Burning the heater of the first vaporizer at low current gave adequate illumination and approximate regulation of the rate of vaporization could be obtained by observation of the rate of deposition on the clear section of bulb thus exposed.

As in the tubes previously described, the measured films were laid down on a target which could be moved into position to receive any given fraction of either metal and then into the measuring position before the electron gun. The targets *TT* were, in this tube, duplicate Pyrex disks 18 mm in diameter, carefully mounted so that one reproduced the position of the other before the gun. A modification of the outgassing procedure applied in our previous work was necessitated by the relatively high volatility of zinc. The preliminary outgassing at 500–550°C was followed, after filling with dried air and introduction of the zinc and barium charges, by baking at 150°C rather than at 350°. Before sealing the tube from the pumps both barium and zinc were melted down two or three times to remove absorbed gases and these fusions were repeated after seal-off until approximately two-thirds of the original charges had been vaporized. The second vaporizers were flashed after each fusion to remove the condensates. The next fractions, amounting to an estimated 10–20 percent of the total charges, were utilized for the measurements and were distilled slowly to the second vaporizers. One of the targets was then coated with a thick mirror-like film of zinc by slow distillation from the second vaporizer and the potential setting for this surface determined. A succession of zinc films was then formed on top of the original zinc surface and each measured to determine the reproducibility and constancy of the potential setting for zinc. In certain experiments designed to test the possible effects of gas contamination, and discussed below, the films were deposited almost instantaneously by raising the voltage applied

to the second vaporizer. The zinc film was then brought before the second barium vaporizer and a succession of barium films laid down and measured. Thick films of zinc and barium were then formed alternately on this same target and the contact potential difference Ba−Zn thus checked repeatedly. The whole series of measurements was then repeated with the second target, starting with a thick film of barium in immediate contact with the glass target. This procedure was designed to show whether or not the work function of either metal was influenced by changing the substrate from glass to the other metal of the pair. No such influence was, however, detected.

Materials

The barium employed in these measurements was taken from the same stock used in our earlier work and previously described. The zinc was the "spectroscopically pure" metal prepared by the New Jersey Zinc Company and kindly furnished to me by Mr. H. M. Cyr, with a spectroscopic analysis showing total impurities under 0.001 percent.

The Contact Difference of Potential Between Barium and Zinc

For all of the fifteen pairs of barium-zinc surfaces measured the contact differences of potential, mean value 1.76 v, were reproducible to less than ±0.02 v. As indicated above, a measurement for a given "pair" of surfaces included in some instances the determination of potential settings for as many as five individual zinc surfaces and as many barium surfaces. Potential settings for the individual films of such a series were reproducible to ±0.01 v for both zinc and barium and the observed contact P.D. between aging surfaces remained constant to within ±0.02 v for periods of measurement extending from 8 seconds after deposition to several days. The measurements were the most satisfactory yet obtained for any pair of metals studied during the course of this program, in which barium has consistently shown a reproducibility limited only by the constancy of the measuring circuit but the second metal, e.g., silver, magnesium, has been less reproducible. The entire series of measurements on barium-zinc included observations with three tubes, one of which was measured after the electrical circuits were rebuilt to improve their manipulation. Results for the three tubes were indistinguishable within the limits of reproducibility for any one tube.

The External Work Function of Zinc

The work function of barium films prepared by the standard technique described above is now known with a high degree of probable accuracy to be 2.52 ev at room temperature. Combination of this value with our result of 1.76±0.02 v for the contact difference of potential between barium and zinc assigns a value of 4.28±0.02 ev to the external work function of zinc surfaces prepared by thermal vaporization and condensation on glass at room temperature. This result is to be compared with such of the earlier photoelectric and contact potential investigations of zinc as have been made under modern vacuum conditions. In Dillon's photoelectric study of polycrystalline and monocrystalline zinc[6] the massive specimens were prepared in vacuum outside of the photoelectric tube, transferred quickly to the tube and subjected to prolonged pumping and irradiation with ultraviolet light. The irradiation produced progressive changes in the initial threshold and, finally, equilibrium values of 3.32 and 3.57 ev, respectively, for polycrystalline zinc and Zn (0001). Dillon tentatively attributed these values to gas-free surfaces but recognized the possibility that they might have been characteristic of gas-coated surfaces stable under the conditions of the experiment. In the later work of DeVoe,[3] also done in Mendenhall's laboratory, the photoelectrically measured surfaces were prepared by distillation in the measuring tube itself under especially rigorous experimental conditions. The observed thresholds were highly reproducible and and were constant over long periods of time. The Fowler threshold was found to be 4.24 eV at a residual pressure of 1.5×10^{-8} mm. This is the only published value in good agreement with the present results.

Since the present measurements were taken in sealed-off tubes it was not possible to test DeVoe's interesting observations: (1) that the

[6] J. H. Dillon, Phys. Rev. **38**, 408 (1931).

work function of zinc remains unchanged as the residual air pressure is increased from 1.5×10^{-8} to 3×10^{-7} mm, but then drops rapidly to 3.5–3.6 ev as the pressure is increased to 10^{-6} mm, and (2) that the work function of zinc is insensitive to pure nitrogen up to at least 10^{-3} mm, the highest pressure tested. These observations are, however, borne out by the measurements of Oatley,[4] who obtained a work function of 3.4 ev at a residual pressure of 10^{-6} mm. An extreme sensitivity of the work function to residual air pressure at pressures in the neighborhood of 10^{-6} mm, such as is indicated by DeVoe's curve, would account also for the work function fluctuations observed by Oatley.

Effects of Gas Contamination at Lowest Residual Pressures

Is the work function given by the present measurements, 4.28 ev, characteristic of clean zinc or must this value be attributed to a contaminated surface in equilibrium with residual gas? In work function determinations by both the photoelectric and contact potential methods it has been customary to utilize the most rigorous vacuum and outgassing techniques possible and then, provided only that the observed values are constant and reproducible, to assign these observed values to clean metal surfaces. But contaminating gas films must be built up to some degree even in the highest attainable vacuum and, after equilibrium between adsorbed film and bulk gas has been established, reproducibility and constancy of work function is to be expected for a contaminated as well as for a clean surface. The fundamental importance of this difficulty is generally recognized but the problem which it presents has received little explicit attention.

The experimental conditions which appear to be prerequisite to an attack upon the problem are: (1) the metal surface must be so prepared as to be initially free from gas, and (2) the first work function measurement must be made before a gas film, sufficiently concentrated to produce a measurable change in work function, has had time to build up. If these conditions are fulfilled and if further measurements show no drift in work function, we may conclude that the equilibrium gas film is either too dilute to affect the work function or is composed of gases to which the work function is insensitive. While the first of these conditions can hardly be realized in its ideal form it is probably approached when the metal surface is prepared by multiple distillation in a gettered vacuum and the final vaporization is not too slow. If the maximum time lapse allowable under condition (2) could be calculated for any given metal-gas system it would then be possible to set up definite criteria for the measurement of any given metal. Such a calculation, involving time rates of formation of adsorbed gas films, evidently can be made only by a kinetic method of the Langmuir type and the minimum data required will not, in general, be available. In certain cases however, as shown below, enough information can be obtained from work function measurements themselves to enable us to get around this difficulty.

In the present measurements no drift in work function could be detected for freshly deposited zinc surfaces within a series of measurements beginning 8 seconds and extending to several days after formation of the surface. To obtain an estimate of the amount and probable effects of the gas deposited in the 8-second interval we compare the behavior of zinc with that of a suitably chosen "test surface" which is initially freed from gas and, like zinc, measured in a tube which has been gettered by vaporized barium, i.e., under a residual gas of comparable pressure and composition. The test surface is so selected that (1) it is sensitive to the same component of the residual gas as is zinc, and (2) its adsorption mechanism for this active gas is already known. The behavior of the test surface then furnishes an index to the rate at which the active gas, of otherwise unknown concentration, is supplied to the zinc surface. Since DeVoe's gas tests show that oxygen is the active contaminant for zinc, tungsten suggests itself as an ideal test surface. The reflection coefficient of tungsten for oxygen is known to be extremely small and the work function of tungsten is extraordinarily sensitive to traces of adsorbed oxygen. Now it has been found in numerous published[7] and unpublished measurements in this laboratory that tungsten, cleaned by flashing to 2800°K

[7] Reference 1; first citation.

and then allowed to cool, invariably shows a drift of work function, but that, in well gettered tubes, this drift does not become measurable until several minutes after flashing. In tubes immersed in liquid air the drift may be postponed for as much as 20 minutes. The drift is toward higher work function and continues for several hours before equilibrium is attained. In the earliest stages of adsorption n, the number of adsorbed molecules per unit surface of adsorber, is too small to affect the time rate of increase in n and we may write

$$dn=(1-\alpha)\nu dt,$$

where α is the reflection coefficient and ν is the number of oxygen molecules which strike unit area in unit time. The effect of the inert components of the residual gas upon α presumably will be of the same order of magnitude for both tungsten and zinc and in the gettered vacuum ν will be approximately the same for both surfaces. In terms of this equation our observations on tungsten show that dn/dt is so small that n attains the extremely small value necessary for detection only after a time interval of the order of minutes. Turning to zinc we may conclude, then, that even if all the oxygen molecules which strike its surface in 8 seconds were to remain adsorbed their number would be insufficient to produce any significant change in work function within this time interval. From the absence of a time drift of the work function of zinc, and its occurrence for tungsten, we may infer, furthermore, that the mean time of sojourn, τ, of an adsorbed oxygen molecule must be much smaller on zinc than on tungsten. For it follows from purely statistical considerations which are independent of adsorption mechanisms, that for any equilibrium film of adsorbed gas

$$\tau=n/\nu.$$

Even when ν is small, τ is so large for oxygen on tungsten that at equilibrium n approaches the total number of adsorbing positions available per unit area. The fact that zinc, although sensitive to oxygen, shows no drift at these small values of ν can be accounted for only if n and τ are so small that the equilibrium film when finally completed is too dilute to produce a measurable change in the work function. We have, then, reasonably convincing evidence that the work function of a zinc surface prepared and measured in a barium gettered vacuum is characteristic of the uncontaminated metal. It may be mentioned, finally, that from the point of view of residual gas effects the contact potential method possesses a unique advantage over emission methods for determining work functions. Since it measures the average work function of a surface, the contact potential is relatively insensitive to slight contaminations which, localized on emitting areas, may have a relatively serious effect upon the emission characteristics of the surface.

The Water Vapor Molecule

BYRON T. DARLING AND DAVID M. DENNISON
University of Michigan, Ann Arbor, Michigan
(Received November 9, 1939)

The problem of the vibration rotation spectrum of water vapor is treated by means of the theory of semi-rigid polyatomic molecules developed by Wilson and Howard. The potential energy is expanded as a power series in the normal coordinates and involves three zeroth-order constants, six first-order and six second-order constants. The positions of the band centers are calculated and found to depend upon ten quantities, X_i, X_{ik}, and γ which are functions of the potential constants. A new feature of the treatment is the recognition of a resonance interaction between certain of the overtone bands which arises from the near equality of ν_1 and ν_3. Eighteen band centers are known experimentally. These serve to determine the X_i, X_{ik}, γ and furnish eight self-consistency checks which are very adequately satisfied. There exists no discrepancy between the Raman and infra-red spectra as reported earlier. In order to obtain the geometric displacements corresponding to each normal coordinate it is necessary to examine the spectrum of D_2O. This not only furnishes the required information but also allows two independent checks upon the theory both of which turn out to be in nearly perfect accord. The interaction between vibration and rotation is considered and the effective moments of inertia are calculated. These are functions of the normal frequencies and of the first-order potential constants. It is shown that $\Delta = I_C - I_A - I_B$ depends only upon the normal frequencies and hence may be computed at once. A comparison between the observed and predicted Δ yields a very satisfactory agreement. The analysis of the rotational structure made by Mecke is supplemented by taking account of the rotational stretching. The resulting molecular constants fix the valence angle to be 104°31′ and the O–H distance to be 0.9580A. From the effective moments of inertia the first-order potential constants may be evaluated and these, together with X_{ik} determine the second-order potential constants. It is now possible to compute the interaction constant γ and a comparison with the observed γ again results most satisfactorily.

1. INTRODUCTION

THE vibration rotation spectrum of water vapor is a very extensive one and has been the subject of a great many investigations. The first strong line of the rotation series must occur in the far infra-red at about 500μ and the spectrum extends with very few gaps through the near infra-red, where the fundamental vibration frequencies lie, up to about the middle of the visible spectrum. It has been accurately mapped from 135μ to 5700A. The far infra-red lines, which correspond to changes in the rotational energy only, have been measured with great precision by Randall, Dennison, Ginsburg and Weber[1] and lead to a determination of the rotational energy levels of the molecule.

It is well known that the water molecule, which has the form of an isosceles triangle, possesses three normal modes of vibration. These fundamental oscillations are shown in Fig. 1. Each of them should be active both in the infra-red and in the Raman spectrum. In point of fact, however, only ν_2 and ν_3 have been observed[2] in the infra-red. They lie near 1595 and 3755 cm^{-1}, respectively. These bands are both extremely intense and the third fundamental ν_1 is probably weaker and is masked by the absorption of ν_3. The principal feature of the Raman spectrum[3] is a strong line at about 3650 cm^{-1} which can be ascribed to ν_1. In addition to the three fundamentals, some fifteen harmonic and overtone bands have been measured. These extend through the near infra-red into the vidible region of the spectrum where they are observed as atmospheric absorption lines in the solar spectrum. The accuracy with which these latter bands have been mapped is very high indeed although not essentially better than that attained in measuring the far infra-red rotation lines.

The fine structure of each band, whether fundamental or overtone, is quite extensive since the moments of inertia of the molecule are small and the rotational energies large. The structure appears to be very irregular, as must be expected since no two of the moments of inertia are equal and the molecule belongs to the class of asym-

[1] H. M. Randall, D. M. Dennison, N. Ginsburg and L. R. Weber, Phys. Rev. **52**, 160 (1937).
[2] E. K. Plyler and W. W. Sleator, Phys. Rev. **37**, 1493 (1931).
[3] D. H. Rank, K. D. Larsen and E. R. Bordner, J. Chem. Phys. **2**, 464 (1934); D. Bender, Phys. Rev. **47**, 252 (1935).

metric rotators. Mecke[4] and his co-workers have made a careful study of the rotational structure of all the known bands and have succeeded in analyzing them in the light of the theory of the asymmetric rotator. The results of their work may be stated as follows. Neglecting a rotational stretching effect which will be discussed later, the rotational levels of the molecule in any particular vibrational state may be organized by means of three moments of inertia. These moments, which will be called effective moments of inertia, vary slightly with the vibrational state in question. Mecke showed that they might be expressed as constants plus small corrections which are linear in the vibrational quantum numbers. By this means he extrapolated to the moments of inertia of the nonvibrating molecule and obtained the molecular dimensions. He found the valence angle to be 104° 36′ and the O–H distance to be 0.9558A. A characteristic feature of the effective moments of inertia is that they do not satisfy the relationship $I_A+I_B=I_C$ which must be expected to hold for any coplanar molecule. In a very important contribution to the theory of semi-rigid rotators Wilson and Howard[5] have shown that the rotational energy levels of a molecule should exhibit just the features observed by Mecke, namely that they may be correlated by means of effective moments of inertia which will vary with the vibrational state.

An analysis of the positions of the centers of the fundamental and overtone bands has been made by Bonner[6] who used the conventional expression for the vibrational energy,

$$W=W_0+\sum X_i n_i+\sum_{i\geq k}\sum X_{ik} n_i n_k$$

where n_1, n_2 and n_3 are the vibrational quantum numbers. The nine constants X_i, X_{ik}, may be determined from the eighteen observed bands and thus allow nine self-consistency checks. These checks are moderately good with one important exception. The fundamental frequency ν_1, measured as a Raman line, is predicted to lie from forty to fifty wave numbers below its observed position.

[4] R. Mecke, W. Baumann and K. Freudenberg, Zeits. f. Physik **81**, 313, 445, 465 (1933).
[5] E. B. Wilson and J. B. Howard, J. Chem. Phys. **4**, 260 (1936).
[6] L. G. Bonner, Phys. Rev. **46**, 458 (1934).

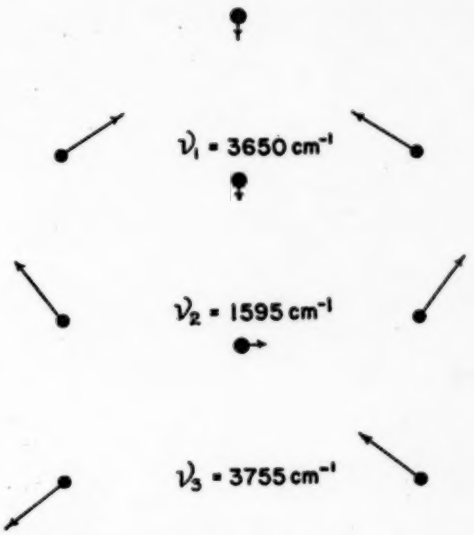

FIG. 1. Fundamental oscillations of the water molecule.

The solution of this difficulty appears to lie in the fact that the two fundamentals ν_1 and ν_3 have nearly the same value, thus paving the way for resonance effects. These effects are not of first order, however, since ν_1 and ν_3 have different symmetry characters and consequently may not interact directly. It is well known that the vibrational energy levels of the nonlinear, symmetrical molecule YX_2 belong to either of two symmetry classes. One of these is even upon reflection in a plane perpendicular to the X–X line and through the Y atom, and it is characterized by n_3 being equal to an even integer. The other class is odd upon reflection and occurs when n_3 is an odd integer. In designating a level it will be convenient to employ the symbol[7] (n_3,n_1,n_2). It is thus obvious that the levels which may interact due to the near equality of ν_1 and ν_3 will be of the type (n_3,n_1,n_2) and (n_3+2,n_1-2,n_2). The hypothesis that a resonance of this sort constitutes an essential feature of the analysis receives immediate confirmation when we examine the overtone bands which occur with greatest intensity. A glance at Table I shows that the strong overtone bands usually appear in pairs, thus (120), (300); (121), (301); (310), (130), etc. The members of these pairs comprise just those levels which could interact by resonance. An analysis of the positions of the

[7] The order (n_3,n_1,n_2) corresponds with the notation used by Mecke which it seems wise to retain here.

TABLE I. *Observed and calculated values of W/hc.*

$(n_3n_1n_2)$	W/hc(obs.)	W/hc(calc.)
(001)	1595.0	—
(002)	3151.0	—
(010)	3650	—
(100)	3755.41	—
(101)	5330.8	5330.6
(110)	7251.0	7250.4
(111)	8807.05	8805.52
(120)	10613.12	—
(300)	11032.36	—
(121)	12151.22	[12148.46
(301)	12565.01	12567.74]
(130)	13830.92	—
(310)	14318.77	—
(131)	15347.91	[15346.27
(311)	15832.47	15834.12]
(132)	16821.61	16822.70
(140)	16899.01	16894.25
(320)	17495.48	17482.60

energy levels will require in addition to the nine X_i and X_{ik} of the usual theory, one further constant γ which measures the effect of the resonance. It will be shown that, upon determining the ten constants with ten of the observed levels, the remaining eight levels are satisfactorily predicted by the theory.

The general program which we hope to carry out in the present paper is as follows: The Hamiltonian function will be set down and the potential developed as a power series in the normal coordinates. The zeroth-order potential contains three constants, namely, the normal frequencies. The first-order potential consists of cubic combinations of the coordinates and contains six constants. The second-order potential which involves quartic combinations possesses also six arbitrary constants. A solution is made which yields the energy values of the system and is a function both of the vibrational and of the rotational quantum numbers. The band centers are obtained by setting the rotational numbers equal to zero and in this way the ten constants X_i, X_{ik} and γ which were previously determined from the observed band centers are found as functions of the fifteen potential constants. To proceed further we turn to the rotational energy of the system and calculate the effective moments of inertia of the molecule. These are found to depend linearly upon the vibrational quantum numbers in agreement with Mecke's observations. They are functions of the six cubic constants but not of the quartic constants. The effective moments of inertia as determined experimentally furnish us with nine data, the coefficients of n_1, n_2 and n_3 for each of the three moments, in addition to the equilibrium values of the moments. These nine data fix the six cubic constants and allow three independent checks of the theory, which are made by comparing the observed and predicted values of $\Delta=I_C-I_A-I_B$. The agreement with the observations is most satisfactory. We now possess sixteen data, the six cubic constants from the effective moments of inertia and the ten X_i, X_{ik} and γ with which to calculate the fifteen potential constants. This implies the existence of one further test of the self-consistency of the theory and upon calculation it proves to be adequately fulfilled.

2. THE HAMILTONIAN

In recent years a number of investigators have attacked the problem of finding the Hamiltonian of a semi-rigid, rotating polyatomic molecule. Of these, the formulation which seems to be most convenient in application is that by Wilson and Howard.[5] A second derivation of this Hamiltonian by a somewhat shorter and more direct method has been made by one of us (B. T. D.) and will be published shortly. The result of this calculation is to obtain a Hamiltonian which differs only slightly and almost unessentially from the Wilson and Howard function, but which is Hermitian while it may be shown that theirs is not.

Our Hamiltonian, when expressed in the Wilson and Howard notation, is:

$$H=\tfrac{1}{2}\sum_{\alpha\beta}\mu^{\frac{1}{2}}(P_\alpha-p_\alpha)\mu_{\alpha\beta}\mu^{-\frac{1}{2}}(P_\beta-p_\beta)\mu^{\frac{1}{2}}+\tfrac{1}{2}\sum_k\mu^{\frac{1}{2}}p_k\mu^{-\frac{1}{2}}p_k\mu^{\frac{1}{2}}+V.$$

The P_α are components of the total angular momentum and p_α are the components of the local angular momentum in the rotating coordinate system defined by Wilson and Howard. $\mu_{\alpha\beta}$ are the cofactors of a determinant μ which contains the moments of inertia. These quantities depend upon the normal coordinates and will be calculated in detail later. p_k are the momenta conjugate to the normal coordinates q_k.

In the case of the water molecule, some simplifications occur. It will be shown that, of the local angular momenta, the only nonvanishing component will be perpendicular to the plane of the molecule. We denote it by p_3. The potential function V we develop as a power series in the normal coordinates where the zeroth-order, first-order and second-order portions will contain terms which are, respectively, quadratic, cubic and quartic in the normal coordinates. Since the molecule is assumed to be symmetrical with respect to a plane perpendicular to the line joining the equilibrium positions of the hydrogen atoms and passing through the oxygen, the potential must obviously be an even function of q_3, where q_3 describes the antisymmetrical motion ν_3 (see Fig. 1). The Hamiltonian, when written out, thus becomes:

$$H=\tfrac{1}{2}\sum_{\alpha=1}^{3}\mu_{\alpha\alpha}P_\alpha^2+\tfrac{1}{2}\mu_{12}(P_1P_2+P_2P_1)+\tfrac{1}{2}\mu^{\frac{1}{2}}p_3\mu_{33}\mu^{-\frac{1}{2}}p_3\mu^{\frac{1}{2}}-\tfrac{1}{2}(p_3\mu_{33}+\mu_{33}p_3)P_3+\tfrac{1}{2}\sum_{k=1}^{3}\mu^{\frac{1}{2}}p_k\mu^{-\frac{1}{2}}p_k\mu^{\frac{1}{2}}$$
$$+\tfrac{1}{2}(\lambda_1q_1^2+\lambda_2q_2^2+\lambda_3q_3^2)+\tfrac{1}{2}(a_1q_1^3+a_2q_2^3+a_3q_1^2q_2+a_4q_1q_2^2+a_5q_1q_3^2+a_6q_2q_3^2)$$
$$+\tfrac{1}{2}(b_1q_1^4+b_2q_2^4+b_3q_3^4+b_4q_1^2q_2^2+b_5q_1^2q_3^2+b_6q_2^2q_3^2). \quad (1)$$

The constants λ_1, λ_2, λ_3, $a_1\cdots a_6$, $b_1\cdots b_6$ constitute the fifteen arbitrary potential constants which were mentioned in the Introduction.[8]

3. The Vibrational Energy

In order to obtain an expression for the band centers we have merely to set the total angular momenta components P_α equal to zero and examine the remainder of the Hamiltonian. It will prove convenient at this point to use, instead of the normal coordinates q_i, the dimensionless variables x_i. The x_i are the arguments of the Hermitian orthogonal functions describing the normal vibrations and are related to q_i as follows: $x_i=2\pi[\omega_ic/h]^{\frac{1}{2}}q_i$. The ω_i appearing here and throughout the remainder of the paper are the normal frequencies expressed in waves per cm; c is the velocity of light. The potential energy in these coordinates becomes:

$$V/hc=\tfrac{1}{2}(\omega_1x_1^2+\omega_2x_2^2+\omega_3x_3^2)+(\alpha_1x_1^3+\alpha_2x_2^3+\alpha_3x_1^2x_2+\alpha_4x_1x_2^2+\alpha_5x_1x_3^2+\alpha_6x_2x_3^2)$$
$$+(\beta_1x_1^4+\beta_2x_2^4+\beta_3x_3^4+\beta_4x_1^2x_2^2+\beta_5x_1^2x_3^2+\beta_6x_2^2x_3^2).$$

The relationship between the α_i, β_i and the a_i, b_i of the earlier formula may be easily obtained when needed.

The solution of this part of the problem which describes the vibrating but nonrotating molecule may be readily obtained by the usual perturbation methods. One finds

$$W/hc=W_0/hc+X_1n_1+X_2n_2+X_3n_3+X_{11}n_1^2+X_{22}n_2^2+X_{33}n_3^2+X_{12}n_1n_2+X_{13}n_1n_3+X_{23}n_2n_3.$$

[8] It should be pointed out that the quartic part of the potential may contain three additional terms, namely $q_1q_2q_3^2$, $q_1q_2^3$ and $q_1^3q_2$. These terms, since they are odd in at least one coordinate, will not contribute to the energy. They would, however, have had a slight influence in transforming from the Hamiltonian for H_2O to that for D_2O, since the normal coordinates of the latter molecule are not exactly equal to those of the former.

The constants X_i and X_{ik} are the following functions of the potential constants:

$$W_0/hc=\tfrac{1}{2}(\omega_1+\omega_2+\omega_3)+\tfrac{1}{4}(X_{11}+X_{22}+X_{33}+X_{12}+X_{13}+X_{23}),$$

$$X_1=\omega_1+X_{11}+\tfrac{1}{2}(X_{12}+X_{13}),$$

$$X_2=\omega_2+X_{22}+\tfrac{1}{2}(X_{12}+X_{23}),$$

$$X_3=\omega_3+X_{33}+\tfrac{1}{2}(X_{13}+X_{23}),$$

$$X_{11}=\text{K.E.P.}+\frac{3}{2}\beta_1-\frac{15}{4}\frac{\alpha_1^2}{\omega_1}-\frac{\alpha_3^2}{2\omega_2}-\frac{\alpha_3^2}{8(2\omega_1+\omega_2)}+\frac{\alpha_3^2}{8(2\omega_1-\omega_2)},$$

$$X_{22}=\text{K.E.P.}+\frac{3}{2}\beta_2-\frac{15\alpha_2^2}{4\omega_2}-\frac{\alpha_4^2}{2\omega_1}-\frac{\alpha_4^2}{8(2\omega_2+\omega_1)}+\frac{\alpha_4^2}{8(2\omega_2-\omega_1)},$$

$$X_{33}=\text{K.E.P.}+\frac{3}{2}\beta_3-\frac{\alpha_5^2}{2\omega_1}-\frac{\alpha_6^2}{2\omega_2}-\frac{\alpha_5^2}{8(2\omega_3+\omega_1)}+\frac{\alpha_5^2}{8(2\omega_3-\omega_1)}-\frac{\alpha_6^2}{8(2\omega_3+\omega_2)}+\frac{\alpha_6^2}{8(2\omega_3-\omega_2)},\tag{2}$$

$$X_{12}=\text{K.E.P.}+\beta_4-\frac{3\alpha_1\alpha_4}{\omega_1}-\frac{3\alpha_2\alpha_3}{\omega_2}-\frac{\alpha_3^2}{2(2\omega_1+\omega_2)}-\frac{\alpha_3^2}{2(2\omega_1-\omega_2)}-\frac{\alpha_4^2}{2(2\omega_2+\omega_1)}-\frac{\alpha_4^2}{2(2\omega_2-\omega_1)},$$

$$X_{13}=\text{K.E.P.}+\beta_5-\frac{3\alpha_1\alpha_5}{\omega_1}-\frac{\alpha_3\alpha_6}{\omega_2}-\frac{\alpha_5^2}{2(2\omega_3+\omega_1)}-\frac{\alpha_5^2}{2(2\omega_3-\omega_1)},$$

$$X_{23}=\text{K.E.P.}+\beta_6-\frac{3\alpha_2\alpha_6}{\omega_2}-\frac{\alpha_4\alpha_5}{\omega_1}-\frac{\alpha_6^2}{2(2\omega_3+\omega_2)}-\frac{\alpha_6^2}{2(2\omega_3-\omega_2)}.$$

A calculation similar to ours was first made by Bonner.[6] Our results for that part of the X_{ik} which depends upon the potential constants is in agreement with his when allowance is made for the difference in notation.

The symbol K.E.P. which appears in the above formulas for the X_{ik} represents the contribution from the kinetic energy part of the Hamiltonian. An examination of the Hamiltonian Eq. (1), shows that the K.E.P. may come from a number of terms, many of which turn out to be quite negligible. Thus the difference between $\frac{1}{2}\sum\mu^{\frac{1}{2}}p_k\mu^{-\frac{1}{2}}p_k\mu^{\frac{1}{2}}$ and the zeroth-order expression $\frac{1}{2}\sum p_k^2$ yields nothing to the order of approximation to which we are working. Similarly $\frac{1}{2}\mu^{\frac{1}{2}}p_3\mu_{33}\mu^{-\frac{1}{2}}p_3\mu^{\frac{1}{2}}$ may be replaced by $\frac{1}{2}\mu_{33}p_3^2$. This latter quantity which involves the single nonvanishing internal angular momentum p_3, will in general contribute both to X_{13} and X_{23}. A calculation shows[9] that for the water molecule the K.E.P. of X_{13} is so small that it may be set equal to zero but that the K.E.P. of X_{23} is $+26.74$ cm^{-1}.

The final step in discussing the formula for the positions of the band centers is to consider the influence of the resonance interaction between levels of the type $(n_3n_1n_2)$ and (n_3+2,n_1-2,n_2). The matrix elements in the Hamiltonian connecting these states are all off diagonal and they are all of second order. Thus the calculation of the X_{ik} will be unaffected, since these depend upon second-order elements on the diagonal but only upon the first-order off-diagonal elements. A second way of stating this is that none of the terms in X_{ik} contain the resonace denominator $(\omega_1-\omega_3)$. The method of deducing the final energy levels is now to diagonalize separately those portions of the Hamiltonian

[9] The calculation is easily made when p_3 is explicitly determined as will be done in the next section. μ_{33} may be given its equilibrium value $1/I_e^0$. The physical reason why X_{13} is so little affected is that in water vapor the motions of the hydrogen atoms in the vibration ν_1 are very nearly along the legs of the isosceles triangle. Thus the compound motion of ν_1 and ν_3 together possesses almost no angular momentum. For this reason the K. E. P. of X_{23} is just equal to the corresponding contribution to X_{23} in CO_2 which has been calculated by Weinberg and Eckart, J. Chem. Phys. **5**, 517 (1937) and found to be $[h/8\pi^2I_e^0][\omega_3/\omega_2+\omega_2/\omega_3]$.

matrix which contain the elements in question. The off-diagonal terms of these portions of the Hamiltonian consist of an interaction b. A short calculation shows that

$$b^{n_3\ n_1\ n_2}_{n_3+2\ n_1-2\ n_2}=\tfrac{1}{2}\gamma hc[n_1(n_1-1)(n_3+1)(n_3+2)]^{\frac{1}{2}},$$

where

$$\gamma=\tfrac{1}{2}\beta_5-\frac{\alpha_5^2}{\omega_1}+\frac{\alpha_1\alpha_5}{2\omega_1}+\frac{\alpha_3\alpha_6}{4(2\omega_1-\omega_2)}-\frac{\alpha_3\alpha_6}{4(2\omega_1+\omega_2)}. \tag{3}$$

It is seen that γ depends only upon the potential constants. There exists a contribution from the kinetic energy part of the Hamiltonian but this turns out to be negligible in the case of the water molecule.

As an example, let us consider the interacting levels (120) and (300). They will be given as the roots of the determinant

$$\begin{vmatrix} X_1+2X_2+X_{11}+4X_{22}+2X_{12}-W/hc & \sqrt{3}\gamma \\ \sqrt{3}\gamma & 3X_1+9X_{11}-W/hc \end{vmatrix}.$$

The roots may be expressed as $\bar{W}/hc\pm\frac{1}{2}[\delta_0^2+\delta_1^2]^{\frac{1}{2}}$ where $\bar{W}/hc$ is the mean value of the levels, in this case, $2X_1+X_2+5X_{11}+2X_{22}+X_{12}$, δ_0 is the separation of the levels without interaction, namely $-2X_1+2X_2-8X_{11}+4X_{22}+2X_{12}$ and δ_1 is the separation caused by the interaction alone, in the present example $2\sqrt{3}\gamma$. The highest frequency bands observed are (140) and (320). These are evidently members of a triplet whose third member is (500) and in this case the determinant which fixes the energy levels is of third order.

The observed positions of the band centers are collected in the Table I. It will be noticed that the numbers given for the bands lying in the near infra-red differ slightly from the values usually appearing in the literature. This arises from the fact that we have corrected the wave-lengths to vacuum.[10] Several measurements[3] have been made of the frequency ν_1 which appears in the Raman spectrum and the values which have been reported range from 3648 cm^{-1} to 3654 cm^{-1}. This discrepancy probably has its origin partly in the broadness of the Raman line (about 5 cm^{-1}). We have adopted the even value 3650 cm^{-1}. In order to calculate the ten constants X_i, X_{ik} and γ, we need ten data and these have been selected as indicated in the third column of Table I. In the case of (121), (301) and (131), (311) the individual levels were not used but only their sums and this is denoted by bracketing their calculated values.

[10] M. Rusch, Ann. d. Physik **70**, 373 (1923).

Since there are eighteen bands which have been observed, there will occur eight checks upon the theory. The agreement between the measured and calculated values is very satisfactory and is definitely much better than that obtained by Bonner. We feel that the small differences which do exist are due to the fact that we are approximating the potential function by a power series development which is broken off after the quartic terms. This view is strengthened by the circumstance that the agreement appears to become poorer for the higher levels. The numerical values of the constants in waves per cm are as follows:

$$\begin{array}{lll} X_1=3693.89 & X_{11}=-43.89 & X_{12}=-\ 20.02 \\ X_2=1614.5 & X_{22}=-19.5 & X_{13}=-155.06 \\ X_3=3801.78 & X_{33}=-46.37 & X_{23}=-\ 19.81 \\ |\gamma|=\ 74.46 & \omega_1=3825.32 & \\ & \omega_2=1653.91 & \\ W_0/hc=4631.25 & \omega_3=3935.59. & \end{array}$$

The zero-point energy W_0 has been calculated on the basis that the energy constant may be developed as a power series in $n_i+\frac{1}{2}$, thus

$$W/hc=\sum\omega_i(n_i+\tfrac{1}{2})+\sum_{i\geq k}\sum X_{ik}(n_i+\tfrac{1}{2})(n_k+\tfrac{1}{2}).$$

We are aware that the conventional second-order perturbation theory does not quite give this result as has been pointed out by King.[11] We believe that this is due to the fact that the perturbation method involves a development

[11] G. W. King, J. Chem. Phys. **5**, 414 (1937).

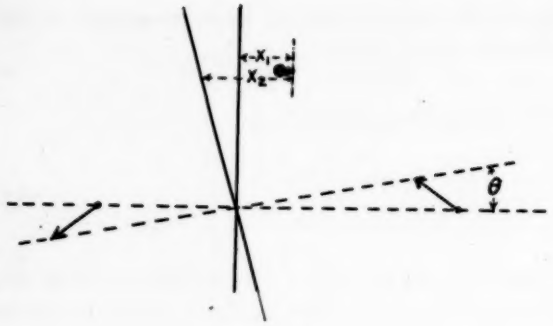

FIG. 2.

which breaks off after a certain number of steps. A similar situation occurs in the theory of the diatomic molecule where the perturbation calculation yields an energy which is not exactly a power series in $(n+\frac{1}{2})$. In this case, however, it is possible to obtain exact solutions for a number of particular potential forms (i.e., the Morse potential) and in every instance the energy is expressed as a function of $n+\frac{1}{2}$.

It will prove necessary for the rotation vibration interaction which is to be treated in the next section that we know the actual geometric displacements corresponding to each of the normal coordinates. The following geometric[12] displacement coordinates will be used. Let us consider the Cartesian axes x and y whose origin lies at the center of gravity of the system and whose x axis is parallel to the line joining the equilibrium positions of the two X atoms. We now choose the x component of the change (increase) of distance between X atoms to be the coordinate q. Let y be the y component of the displacement of the Y atom with respect to the center of gravity of the X atoms. The final coordinate x is illustrated in Fig. 2 and is equal to $x_1+b\theta$, where x_1 is the x component of the displacement of the Y atom against the center of gravity of the X atoms, θ is the angle between the line joining the X atoms and the x axis, while b is the equilibrium distance between the Y atom and the line joining the X atoms. From the conservation of angular momentum it is easily shown that $x=(1+\mu b^2/2ma^2)x_1$. Here μ is the reduced mass $2mM/(2m+M)$ and $2a$ is the equilibrium distance between the X atoms.

The kinetic energy of the vibrating system may be calculated in the usual manner and is

$$2T=\left[\mu\Big/\left(1+\frac{\mu b^2}{2ma^2}\right)\right]\dot{x}^2+\mu\dot{y}^2+\frac{m}{2}\dot{q}^2.$$

The general potential function will contain only four constants (since it must be even in the coordinate x) and may be written

$$2V=ax^2+by^2+cq^2+2dyq.$$

The normal frequencies are given in terms of the roots $(\lambda_i=4\pi^2\omega_i^2c^2)$ of a determinant and may be expressed as follows:

$$\lambda_1\lambda_2=2(bc-d^2)/m\mu,$$
$$\lambda_1+\lambda_2=b/\mu+2c/m,$$
$$\lambda_3=a(1+\mu b^2/2ma^2)/\mu.$$

The first minors of the determinant yield the relationships between the displacement coordinates and the normal coordinates q_1, q_2, q_3.

$$q=\delta_{11}q_1+\delta_{12}q_2,$$
$$y=\delta_{21}q_1+\delta_{22}q_2,$$
$$x=\delta_{33}q_3,$$

where

$$\delta_{1i}=d/[(\tfrac{1}{2}m\lambda_i-c)^2\mu+\tfrac{1}{2}md^2]^{\frac{1}{2}},$$
$$\delta_{2i}=(\tfrac{1}{2}m\lambda_i-c)/[(\tfrac{1}{2}m\lambda_i-c)^2\mu+\tfrac{1}{2}md^2]^{\frac{1}{2}}, \quad (4)$$

where the index $i=1, 2$.

$$\delta_{33}=[1/\mu+b^2/2ma^2]^{\frac{1}{2}}.$$

The following relations exist between the δ_{ik},

$$\mu\delta_{21}^2+\tfrac{1}{2}m\delta_{11}^2=1, \qquad \delta_{11}^2+\delta_{12}^2=2/m,$$
$$\mu\delta_{22}^2+\tfrac{1}{2}m\delta_{12}^2=1, \qquad \delta_{21}^2+\delta_{22}^2=1/\mu,$$
$$2\mu\delta_{21}\delta_{22}+m\delta_{11}\delta_{12}=0, \qquad \mu\delta_{21}^2-\tfrac{1}{2}m\delta_{12}^2=0,$$
$$\delta_{21}\delta_{11}+\delta_{22}\delta_{12}=0, \qquad \mu\delta_{22}^2-\tfrac{1}{2}m\delta_{11}^2=0.$$

It is clear that the three normal frequencies of H_2O will not serve to determine these relationships, since there are four potential constants involved. The normal frequencies of D_2O will supply the remaining datum.

The fundamental frequencies of D_2O have been measured[13] although with somewhat less accu-

[12] A geometric coordinate is one which defines the configuration of the molecule without reference to the masses of the atoms. The constants appearing in the potential function are thus the same for any isotopic molecule.

[13] The frequency ν_1 appears in the Raman spectrum and was reported by D. H. Rank, K. D. Larsen and E. R. Bordner, J. Chem. Phys. 2, 464 (1934). The bands ν_2 and ν_3 were observed by E. F. Barker and W. W. Sleator, J.

racy than the corresponding work for H_2O and have the values $\nu_1=2666$, $\nu_2=1178.7$, $\nu_3=2789$ cm^{-1}. These frequencies are not the normal frequencies ω, but, as may be seen from Eq. (2), are connected with them by the formulas, $\omega_1=\nu_1-2X_{11}-\frac{1}{2}X_{12}-\frac{1}{2}X_{13}$, etc. We shall treat the problem of obtaining the X_{ik} by a method similar to that used by Bonner. In the diatomic molecule the coefficient of n^2 varies as the square of the frequency when changing from one isotopic molecule to another. We shall accordingly assume that X_{ii} varies as ω_i^2 and X_{ik} as $\omega_i\omega_k$. It will turn out that all three normal frequencies change by about the same factor and consequently we feel that this method, while not rigorously accurate, will lead to no serious errors. A preliminary solution gives the variation of each ω. We find in this way that $\omega_1=2757.86$ and $\omega_2=1210.17$.These numbers, which are derived from the observed fundamentals, may be tested by noticing that the ratio of $\omega_1\omega_2$ for H_2O to $\omega_1\omega_2$ for D_2O must be equal to the square root of the ratio of $m\mu$ for D_2O and H_2O, respectively. A calculation based on H=1.00813, D=2.01473, and O=16.0000 shows that this relation is almost perfectly fulfilled; the observed ratio being 1.89566, while the predicted ratio is 1.89539. Such an excellent agreement implies that our method of obtaining the X_{ik} for D_2O has been a valid one. For purposes of computation ω_1 is increased by 0.20 cm^{-1} and ω_2 by 0.08 cm^{-1} in order that the theoretical ratio of $\omega_1\omega_2$ for H_2O and D_2O shall be exactly satisfied.

The third fundamental will be treated slightly differently: ω_3 of D_2O may be obtained by using Bonner's values for the molecular dimensions, $a=0.75625$A and $b=0.58450$A, together with the ratio of masses. We find $\omega_3=2883.79$. From this, together with the X_{ik} we calculate that $\nu_3=2787.7$, which is in very satisfactory accord with the observed number 2789.

Collecting these results, we have for D_2O:

$$\begin{aligned}&\omega_1=2758.06, \quad X_{11}=-22.81, \quad X_{12}=-10.56,\\ &\omega_2=1210.25, \quad X_{22}=-10.44, \quad X_{13}=-81.92,\\ &\omega_3=2883.79, \quad X_{33}=-24.90, \quad X_{23}=-10.62.\end{aligned}$$

The zero-point energy of D_2O is then 3385.74 cm^{-1}.

The values of the potential constants and of the δ_{ik} for H_2O now follow readily.

$$\begin{aligned}&a=10.672\times10^5 \text{ dynes/cm} \quad c=3.1344,\\ &b=7.1810, \qquad\qquad\quad d=3.1588,\\ &\delta_{11}=+1.14678m^{-\frac{1}{2}}, \quad \delta_{12}=+0.82758m^{-\frac{1}{2}},\\ &\delta_{21}=+0.43910m^{-\frac{1}{2}}, \quad \delta_{22}=-0.60845m^{-\frac{1}{2}},\\ &\delta_{33}=\ \ 0.92828m^{-\frac{1}{2}},\end{aligned}$$

where m is the mass of a hydrogen atom, which we will take to be 1.6734×10^{-24} g.

The sign of d is not determined directly but if we assume that the force field must be nearly of the valence type it follows that d is positive. We wish to remark that the accuracy with which the potential constants are known is certainly not as high as would be indicated by the number of significant figures given above. In making calculations, however, it is wise to make sure that no inconsistencies arise from rounding off numbers prematurely.

4. The Vibration Rotation Interaction

The main objective of the present section is to calculate the effective moments of inertia of the molecule. In a very recent paper Shaffer and Nielsen[14] have also undertaken to solve this problem for the nonlinear symmetrical molecule YX_2. They begin, as we do, with the Wilson and Howard Hamiltonian and carry through their computation along lines which in a general way parallel our work. Their final results for the moments of inertia are partly in agreement with ours but contain certain differences which appear to us to be vitally important in a discussion of the problem. Thus we find that the quantity $\Delta=I_C-I_A-I_B$ is independent of the coefficients of the cubic terms in the potential, a result which

Chem. Phys. **3**, 660 (1935). Of these two frequencies ν_3 is the less accurately known and was determined as the center of the zero branch associated with the band. We believe that the published value of 2784 cm^{-1} is subject to the following corrections. In H_2O, the center of the zero branch of ν_3 lies at about 3741 cm^{-1} whereas the true band center as revealed by the rotational analysis is at 3755.4 cm^{-1}. This difference of 14.4 cm^{-1} is caused by a type of convergence. Now in the case of the diatomic molecule the convergence factor varies as the cube of the normal frequency and consequently we believe that the correction for the frequency ν_3 of D_2O should be approximately 5.7 cm^{-1}. A reduction to vacuum wavelengths yields the final result, $\nu_3=2789$ cm^{-1}. Fortunately the frequency ν_3 is not involved in the calculation of the potential constants, although it does furnish a test of the self-consistency of the theory.

[14] W. H. Shaffer and H. H. Nielsen, Phys. Rev. **56**, 188 (1939).

is not attained by Shaffer and Nielsen.* We feel justified in giving a brief account of our calculations, particularly since they appear to be very simple and straightforward.

The moments and production of inertia in the rotating coordinate system defined by Wilson and Howard may be readily expressed as functions of the displacement coordinates x, y and q of the last section. Letting $1/f=(1+\mu b^2/2ma^2)$, we find:

$$A=\mu b^2+2\mu by+\mu^2f^2b^2x^2/2ma^2+\mu y^2,$$
$$B=2ma^2+2maq+\tfrac{1}{2}mq^2+\mu f^2x^2,$$
$$C=A+B,$$
$$D=2\mu bfx+\tfrac{1}{2}\mu fx(bq/a+2y),$$
$$E=F=0.$$

The $\mu_{\alpha\beta}$ which appear in the Hamiltonian function are

$$\mu_{11}=\frac{B}{AB-D^2},\quad \mu_{22}=\frac{A}{AB-D^2},\quad \mu_{12}=\frac{D}{AB-D^2},$$

$$\mu_{33}=\frac{1}{C-\Sigma Z_k^2},\quad \mu_{13}=\mu_{23}=0.$$

The Z_k are obtained in the following manner. The only component of the internal angular momenta p_α which does not vanish is p_3 and has the value

$$p_3=\mu f(x\dot{y}-\dot{x}y)-(\mu fb/2a)(x\dot{q}-\dot{x}q).$$

This may be expressed as a function of the normal coordinates by means of Eq. (3).

$$p_3=K_1(q_3\dot{q}_1-q_1\dot{q}_3)+K_2(q_3\dot{q}_2-q_2\dot{q}_3),$$

where

$$K_1=\mu f\delta_{33}(\delta_{21}-b\delta_{11}/2a),$$
$$K_2=\mu f\delta_{33}(\delta_{22}-b\delta_{12}/2a).$$

It may readily be shown from the relations connecting the $\delta_{\alpha\beta}$ that $K_1^2+K_2^2=1$. A substitution yields the result that for H_2O,[15] $K_1^2=0.000019$ and $K_2^2=0.999981$. The quantity Z_k is the coefficient of $\dot{q}_k$ in the formula for p_3. Thus

$$Z_1=K_1q_3,\quad Z_2=K_2q_3,\quad Z_3=-K_1q_1-K_2q_2.$$

The $\mu_{\alpha\beta}$ may now be expressed as functions of the normal coordinates. It is convenient to develop them as power series in the q_i and to retain no terms beyond the second.

$$\mu_{11}=1/\mu b^2-2(\delta_{21}q_1+\delta_{22}q_2)/\mu b^3+3f^2\delta_{33}^2q_3^2/2ma^2b^2+3(\delta_{21}q_1+\delta_{22}q_2)^2/\mu b^4,$$

$$\mu_{22}=1/2ma^2-(\delta_{11}q_1+\delta_{12}q_2)/2ma^3+3\mu f^2\delta_{33}^2q_3^2/4m^2a^4+3(\delta_{11}q_1+\delta_{12}q_2)^2/8ma^4,$$

$$\mu_{33}=\mu f^2\delta_{33}^2/2ma^2-\frac{f^4\mu^2\delta_{33}^4}{2m^2a^4}[(ma\delta_{11}+\mu b\delta_{21})q_1+(ma\delta_{12}+\mu b\delta_{22})q_2]+3f^4\mu^2\delta_{33}^4(K_2^2q_1^2+K_1^2q_2^2)/4m^2a^4$$
$$-f^4\mu^2\delta_{33}^4K_1K_2q_1q_2/2m^2a^4,$$

$$\mu_{12}=f\delta_{33}q_3/ma^2b-3f\delta_{33}q_3(\delta_{21}q_1+\delta_{22}q_2)/4ma^2b^2-3f\delta_{33}q_3(\delta_{11}q_1+\delta_{12}q_2)/4ma^3b.$$

A feature of these expressions which will later turn out to be important is that the linear part of $1/\mu_{33}$ is equal to the linear part of $(1/\mu_{11}+1/\mu_{22})$. This arises directly from the fact that $A+B=C$.

The Hamiltonian (Eq. (1)) may now be re-examined and each term evaluated to the desired order of approximation. The method which is employed is to diagonalize the Hamiltonian with respect to the vibrational part of the matrix elements. The resulting expression may then—in agreement with Wilson and Howard's predictions—be put in the form

$$W=P_1^2/2I_A+P_2^2/2I_B+P_3^2/2I_C+\text{vibrational energy}.$$

The I_A, I_B and I_C are the effective amounts of inertia and are functions of the vibrational quantum

* *Note added in proof.*—A re-examination of the article by Shaffer and Nielsen leads us to believe that the differences may well be due to several typographical errors and that our results are essentially in agreement.

[15] The physical meaning of this very unequal division is interesting. K_1 and K_2 represent the contributions to the internal angular momentum from the combined motion ν_1, ν_3 and ν_2, ν_3, respectively. In H_2O the motion of the hydrogen atoms (whose amplitude is so much larger than that of the oxygen) is nearly along the legs of the isosceles triangle in both the motions ν_1 and ν_3 but nearly perpendicular to the legs in ν_2. For this reason the combined motion ν_1, ν_3 will possess practically no internal angular momentum as compared with that of ν_2, ν_3.

numbers. It must be stated that the above equation is true only to second-order approximation; higher orders will bring in quartic combinations of the P_α.

The first part $\frac{1}{2}\sum\mu_{\alpha\alpha}P_\alpha^2$ contains, in addition to the constant $\mu_{\alpha\alpha}^0=1/I_\alpha^0$, terms which are linear in the q_s and terms which are quadratic. These latter are immediately diagonalized, the q_i^2 yielding $h(n_i+\frac{1}{2})/4\pi^2\omega_i c$ while q_iq_k gives no contribution. The linear part of $\mu_{\alpha\alpha}$ may be treated in either of two ways which lead to the same results. The first method recognizes that the presence of a linear term means that the centrifugal forces of rotation have displaced the equilibrium positions of the atoms. We may shift to the new origin by balancing the centrifugal force against the Hookes' law force. This process will eliminate the linear part of $\mu_{\alpha\alpha}$ and will bring in contributions from the cubic part of the potential. The second method is that of the standard second-order perturbation theory; the terms linear in the q_i and those which are cubic in the q_i combine to produce an addition to the energy which may be readily calculated and which is equal to that obtained above.

The second quantity in the Hamiltonian, $\frac{1}{2}\mu_{12}(P_1P_2+P_2P_1)$, yields nothing in this order of approximation. It may likewise be shown that the final term

$$-\tfrac{1}{2}(p_3\mu_{33}+\mu_{33}p_3)P_3$$

can be replaced by the simpler expression $-\mu_{33}^0p_3P_3$.

Obviously $-\mu_{33}^0p_3P_3$ contains no elements which are diagonal in the vibrational quantum numbers and therefore no first-order contribution to the energy. In second order it does, however, give an important term which may be calculated from the usual form

$$(\mu_{33}^0)^2P_3^2\sum_k{}'p_{3k}{}^ip_{3i}{}^k/(W_i^0-W_k^0).$$

It should be mentioned that this term does contain a matrix element connecting the resonating states $(n_3n_1n_2)$ and (n_3+2,n_1-2,n_2) but that its magnitude for H_2O is negligibly small and need not be considered.

The methods which have just been outlined allow us to obtain the effective $\mu_{\alpha\alpha}$. The effective moments of inertia are found by inverting the $\mu_{\alpha\alpha}$ and are given by the following expressions. We present them as functions of the potential constants $\alpha_1\cdots\alpha_6$ rather than $a_1\cdots a_6$. Let $s^2=h/\pi^2b^2c$ and $t^2=h/4\pi^2a^2c$.

$$I_A=I_A^0[1-s(3s\delta_{21}{}^2/4\omega_1+3\delta_{21}\alpha_1/\omega_1{}^{\frac{1}{2}}+\delta_{22}\alpha_3/\omega_2{}^{\frac{1}{2}})(n_1+\tfrac{1}{2})-s(3s\delta_{22}{}^2/4\omega_2+3\delta_{22}\alpha_2/\omega_2{}^{\frac{1}{2}}+\delta_{21}\alpha_4/\omega_1{}^{\frac{1}{2}})(n_2+\tfrac{1}{2})$$
$$-s(3s\mu b^2f^2\delta_{33}{}^2/8ma^2\omega_3+\delta_{21}\alpha_5/\omega_1{}^{\frac{1}{2}}+\delta_{22}\alpha_6/\omega_2{}^{\frac{1}{2}})(n_3+\tfrac{1}{2})],$$

$$I_B=I_B^0[1-t(3t\delta_{11}{}^2/4\omega_1+3\delta_{11}\alpha_1/\omega_1{}^{\frac{1}{2}}+\delta_{12}\alpha_3/\omega_2{}^{\frac{1}{2}})(n_1+\tfrac{1}{2})-t(3t\delta_{12}{}^2/4\omega_2+3\delta_{12}\alpha_2/\omega_2{}^{\frac{1}{2}}+\delta_{11}\alpha_4/\omega_1{}^{\frac{1}{2}})(n_2+\tfrac{1}{2})$$
$$-t(3t\mu f^2\delta_{33}{}^2/2m\omega_3+\delta_{11}\alpha_5/\omega_1{}^{\frac{1}{2}}+\delta_{12}\alpha_6/\omega_2{}^{\frac{1}{2}})(n_3+\tfrac{1}{2})],$$

$$\Delta=I_C-I_A-I_B=[hK_1{}^2\omega_3{}^2/\pi^2c\omega_1(\omega_3{}^2-\omega_1{}^2)](n_1+\tfrac{1}{2})+[hK_2{}^2\omega_3{}^2/\pi^2c\omega_2(\omega_3{}^2-\omega_2{}^2)](n_2+\tfrac{1}{2})$$
$$-[(hK_1{}^2\omega_1{}^2/\pi^2c\omega_3(\omega_3{}^2-\omega_1{}^2))+(hK_2{}^2\omega_2{}^2/\pi^2c\omega_3(\omega_3{}^2-\omega_2{}^2))](n_3+\tfrac{1}{2}).$$

These formulas differ from those of Nielsen and Shaffer only in the parts containing the cubic constants $\alpha_1\cdots\alpha_6$. Our results, in contrast to theirs, show that $\Delta=I_C-I_A-I_B$ is independent of the cubic constants. The reason why this comes about is easily understood. The cubic coefficients in the effective moments of inertia arise only from the terms in $\mu_{\alpha\beta}$ which are linear in the q_i. These terms, as has been pointed out, satisfy the relation that $1/\mu_{33}=1/\mu_{11}+1/\mu_{22}$ and consequently they can contribute nothing to the Δ.

The quantity Δ plays an important role in the theory of the water vapor spectrum since it may be calculated from a knowledge of the normal frequencies and the shape of the molecule only and hence compared directly with the experimental data. Explicitly,

$$\Delta=[0.1644(n_2+\tfrac{1}{2})-0.0122(n_3+\tfrac{1}{2})]\times10^{-40}. \tag{5}$$

The coefficient of $(n_1+\frac{1}{2})$ in the expression for Δ is practically zero since the numerical value of K_1^2 is so small.

It will prove convenient to substitute the numerical values for the normal frequencies of the water molecule into the formulas for the effective moments of inertia in order to facilitate a correlation with the experimental data. We thus obtain

$$\begin{aligned}(I_A-I_A^0)\times 10^{40}&=(-0.0151-0.0001127\alpha_1+0.0001831\alpha_3)(n_1+\tfrac{1}{2})+(-0.0672+0.0005493\alpha_2\\&\quad-0.0000376\alpha_4)(n_2+\tfrac{1}{2})+(-0.0149-0.0000376\alpha_5+0.0001831\alpha_6)(n_3+\tfrac{1}{2}),\\(I_B-I_B^0)\times 10^{40}&=(-0.0290-0.0002136\alpha_1-0.0001807\alpha_3)(n_1+\tfrac{1}{2})+(-0.0349-0.0005421\alpha_2\\&\quad-0.0000712\alpha_4)(n_2+\tfrac{1}{2})+(-0.0280-0.0000712\alpha_5-0.0001807\alpha_6)(n_3+\tfrac{1}{2}).\end{aligned}\tag{6}$$

In the case of the water vapor molecule in which there exists a resonance between the states (n_3,n_1,n_2) and (n_3+2,n_1-2,n_2) there will occur a mixing of the effective moments of inertia for these states. If I_a' and I_a^2 represent the moments of inertia as calculated from Eq. (6) and if I_a^I and I_a^{II} are the actual moments of the resonating states, then it may readily be shown, that

$$\begin{aligned}I_a^{\mathrm{I}}&=I_a'(\delta+\delta_0)/2\delta+I_a^2(\delta-\delta_0)/2\delta,\\I_a^{\mathrm{II}}&=I_a'(\delta-\delta_0)/2\delta+I_a^2(\delta+\delta_0)/2\delta,\end{aligned}\tag{7}$$

where in these expressions δ_0 stands for the separation which would exist between the vibrational levels if there were no resonance interaction and $\delta=[\delta_0^2+\delta_1^2]^{\frac{1}{2}}$ is the actual separation between the levels.

5. Correlation with Experiment

The final problem of this investigation will be to compare the experimental fine structure of the water bands with the theoretical formulas for the effective moments of inertia and thereby to obtain the molecular potential constants. Mecke[4] has analyzed the fine structure lines of each band and has shown that the rotational energy levels are just those which are predicted by the theory of the asymmetrical rotator. He thus calculates the effective moments of inertia for every particular vibrational state. We believe that his values may be somewhat improved by taking into account the rotational stretching of the molecule in the manner proposed by Wilson.[16] It was stated in the last section that the application of a higher order perturbation to the Hamiltonian function would yield terms which are quartic in the components of total angular momenta P_a. The coefficients are functions of the normal frequencies only and may be easily evaluated. The final results of our calculation as given in Table II for a limited number of rotational states show that each level will be decreased by what amounts to a small correction.[17]

Table II. *Calculated values of* $-\delta W$.

J	2_1	2_0	2_{-1}	1_1	1_0	1_{-1}
$-\delta W(\mathrm{cm}^{-1})$	0.331	0.032	0.045	0.015	0.006	0.047

The effective moments of inertia are then obtained from the relations

$$\begin{aligned}h/2\pi^2cI_A&=(2_1-1_{-1})_{\mathrm{exp}}+0.28,\\h/2\pi^2cI_B&=(2_0-1_0)_{\mathrm{exp}}+0.03,\\h/2\pi^2cI_C&=(2_{-1}-1_1)_{\mathrm{exp}}+0.03\end{aligned}$$

and are shown in the second, fourth and sixth columns of Table III in units of 10^{-40} g cm^2. (We take $h=6.624\times10^{-27}$.)*

The observed values of $\Delta=I_C-I_A-I_B$ are given in the seventh column, next to which appear the Δ as calculated from Eq. (5). The agreement is very satisfactory and is best for those levels which lie in regions most favorable for experimental determination. Thus the rota-

[16] E. B. Wilson, J. Chem. Phys. **4**, 526 (1936).

[17] The work of Randall, Dennison, Ginsburg and Weber indicates that for the high levels the rotational stretching may be very large (for the state 11_{11} it is about 280 cm^{-1}); in fact, so large that it could not be simply computed.

* *Note added in proof.*—The physical constants appearing in this paper are taken from a private communication kindly furnished by Professor R. T. Birge.

tional energy levels of the ground state (000) are undoubtedly known with the greatest accuracy. The strong bands (120) to (310) which appear as atmospheric lines in the solar spectrum have also been measured with high precision. The weaker solar bands, as well as the near infra-red bands, have not been as exactly determined. We regard the agreement between Δobs. and Δcalc. as being a real and substantial test of the theory and as constituting one of the more important points of this investigation.

The theory also predicts that the effective moments of inertia should be linear functions of the vibrational quantum numbers. A study of the data reveals that this is essentially the case, although small discrepancies exist which are usually of the order of the experimental errors. We have made the following treatment: The seven most accurately observed bands (000) and (120) to (310) are selected and their moments of inertia are fitted to a linear formula by means of a least-square solution which is subject to two supplementary conditions; first that $I_C^0 = I_A^0 + I_B^0$ and second that the Δ is just equal to the calculated Δ. Of course here, as in every case involving the resonating levels, it is necessary to take account of the mixing of the effective moments of inertia caused by the resonance. This is easily done by means of Eqs. (7) and we arrive at the result,

$$I_A \times 10^{40} = 1.0229 + (0.0213)(n_1 + \tfrac{1}{2}) - (0.1010)(n_2 + \tfrac{1}{2}) + (0.0486)(n_3 + \tfrac{1}{2}),$$

$$I_B \times 10^{40} = 1.9207 + (0.0398)(n_1 + \tfrac{1}{2}) - (0.0249)(n_2 + \tfrac{1}{2}) + (0.0077)(n_3 + \tfrac{1}{2}),$$

$$I_C \times 10^{40} = 2.9436 + (0.0611)(n_1 + \tfrac{1}{2}) + (0.0385)(n_2 + \tfrac{1}{2}) + (0.0441)(n_3 + \tfrac{1}{2}).$$

The values for the effective moments of inertia as calculated from the formula just written are given in the third, fifth and seventh columns of Table III. The agreement with the observed values is satisfactorily good and, as in the case of the Δ, is best for those bands which are most accurately known experimentally, thus leading to the belief that the discrepancies may be due to observational errors.

The equilibrium values for the moments of inertia give the O–H distance to be 0.9580A and the valence angle to be 104° 31′; figures which differ only slightly from those obtained by Mecke[4] and Bonner.[6] A comparison of these formulas with Eq. (6) allows us to determine the six cubic constants. We find

$$\alpha_1 = -322 \text{ cm}^{-1}, \quad \alpha_3 = +1, \quad \alpha_5 = -909,$$
$$\alpha_2 = -47, \quad \alpha_4 = +216, \quad \alpha_6 = +160.$$

Knowing the cubic constants together with the numerical values of the six X_{ik}, it is of course possible to obtain the six quartic constants.

$$\beta_1 = +39 \text{ cm}^{-1}, \quad \beta_3 = +35, \quad \beta_5 = +212,$$
$$\beta_2 = +2, \quad \beta_4 = -116, \quad \beta_6 = -108.$$

It is interesting to notice that the perturbing terms in the potential which connect the levels n_1 and n_3, namely $x_1 x_3^2$ and $x_1^2 x_3^2$, appear to be unusually large. This no doubt accounts both for the resonance between the levels $(n_3 n_1 n_2)$, (n_3+2, n_1-2, n_2) and for the fact that they appear to be the most intense overtones in the spectrum.

Since we now possess all twelve of the anharmonic potential constants, together with the normal frequencies, we may calculate the interaction constant γ by means of Eq. (3). We find the value $\gamma_{calc} = -71.5 \text{ cm}^{-1}$ which may be compared with the experimental figure of $|\gamma_{obs}| = 74.46 \text{ cm}^{-1}$. The agreement, within about 4 percent, is remarkably good and argues strongly for the essential correctness of our interpretation of the water vapor spectrum.

TABLE III. *Observed and calculated values of the effective moments of inertia and of* Δ.

$(n_3 n_1 n_2)$	I_A obs.	I_A calc.	I_B obs.	I_B calc.	I_C obs.	I_C calc.	Δ obs.	Δ calc.
(000)	1.007	1.007	1.929	1.932	3.013	3.016	0.077	0.076
(001)	0.904	0.906	1.899	1.907	3.107	3.055	.304	.241
(100)	1.057	1.056	1.928	1.939	3.048	3.059	.063	.064
(101)	0.988	0.955	1.923	1.915	3.133	3.099	.222	.228
(110)	1.000	1.076	1.995	1.979	3.117	3.120	.122	.064
(111)	0.977	0.976	1.896	1.954	3.200	3.160	.327	.228
(120)	1.110	1.105	2.014	2.012	3.185	3.178	.061	.061
(300)	1.141	1.147	1.969	1.962	3.155	3.151	.045	.042
(121)	1.000	1.003	1.978	1.987	3.240	3.217	.262	.226
(301)	1.046	1.046	1.940	1.947	3.201	3.190	.215	.207
(130)	1.133	1.136	2.040	2.040	3.244	3.233	.071	.057
(310)	1.161	1.158	2.006	2.014	3.217	3.219	.050	.047
(131)	1.030	1.034	2.003	2.015	3.287	3.272	.254	.221
(132)	0.961	0.935	1.998	1.990	3.352	3.310	.393	.386
(140)	1.107	1.224	2.056	2.064	3.283	3.286	.120	−.004
(320)	1.183	1.231	2.006	2.055	3.365	3.282	.176	−.006

Magnesium-Like Spectra of Chlorine, Potassium, Calcium and Scandium

W. L. PARKER AND L. W. PHILLIPS*
Department of Physics, University of Illinois, Urbana, Illinois

(Received November 1, 1939)

About eighty lines belonging to the triplet systems of the magnesium-like ions, Cl VI, K VIII, Ca IX, and Sc X, have been identified. Term values for each ion are reported with respect to the $3s3p\ ^3P_0$ level.

IN two previous papers from this laboratory studies of the sodium-like spectra of Cl[1] and K, Ca, and Sc[2] were reported. It is the purpose of this paper to present data on the magnesium-like spectra of these same elements.

The spectra of the first five ions of the magnesium sequence have been known for some time.[3, 4] Bowen and Millikan[5] reported the wavelengths of the pp' group in Cl VI without attempting to make J value assignments. More recently Edlén[6] has reported on the magnesium-like spectra of the ions Ti XI to Co XVI. The data reported here serve to bridge the gap in the sequence from Cl VI to Sc X with the exception of A VII. A preliminary attempt in this laboratory at the excitation of the intermediate stages of ionization of argon has shown some indications of success, and it is hoped that these data will be available soon.

Classified lines in the spectra of Cl, K, Ca, and Sc are listed in Tables I–IV. In all cases the position of the multiplets was first predicted by means of a constant-first-difference or constant-second-difference graph, and the separations were checked by the $(\Delta\nu)^{\frac{1}{2}}$ law. Since the $3s3p\ ^3P$ term in Cl VI has been identified by its combina-

TABLE I. *Classified lines of* Cl VI.

INT	λ(VAC.)	ν(CM^{-1})	CLASSIFICATION
3*	736.762	135,729	$3s3p\ ^3P_2-3p^2\ ^3P_1$
3*	733.891	136,260	$^3P_1\quad ^3P_0$
4*	730.311	136,928	$^3P_1\quad ^3P_1$
			$^3P_2\quad ^3P_2$
3*	727.537	137,450	$^3P_0\quad ^3P_1$
3*	724.129	138,097	$^3P_1\quad ^3P_2$
3	555.580	179,992	$3s3p\ ^3P_2-3s3d\ ^3D_2$
20	555.485	180,023	$^3P_2\quad ^3D_3$
2	552.053	181,142	$^3P_1\quad ^3D_1$
10	551.992	181,162	$^3P_1\quad ^3D_2$
5	550.355	181,701	$^3P_0\quad ^3D_1$
8	399.995	250,003	$3s3d\ ^3D_3-3s4f\ ^3F_4$
7	399.957	250,027	$^3D_2\quad ^3F_3$
5	399.938	250,039	$^3D_1\quad ^3F_2$
25	325.161	307,540	$3s3p\ ^3P_2-3s4s\ ^3S_1$
20	323.936	308,703	$^3P_1\quad ^3S_1$
15	323.356	309,257	$^3P_0\quad ^3S_1$
3	243.883	410,033	$3s3p\ ^3P_2-3s4d\ ^3D_2$
12	243.854	410,082	$^3P_2\quad ^3D_3$
2	243.208	411,170	$^3P_1\quad ^3D_1$
8	243.194	411,194	$^3P_1\quad ^3D_2$
5	242.885	411,718	$^3P_0\quad ^3D_1$
3	195.227	512,224	$3s3p\ ^3P_2-3s5d\ ^3D_3$
2	194.796	513,358	$^3P_1\quad ^3D_2$

* Intensity as reported by Bowen and Millikan.

TABLE II. *Classified lines of* K VIII.

INT	λ(VAC.)	ν(CM^{-1})	CLASSIFICATION
5	572.790	174,584	$3s3p\ ^3P_2-3p^2\ ^3P_1$
4	569.479	175,599	$^3P_1\quad ^3P_0$
6	565.112	176,956	$^3P_1\quad ^3P_1$
12	564.462	177,160	$^3P_2\quad ^3P_2$
6	561.593	178,065	$^3P_0\quad ^3P_1$
2	557.029	179,524	$^3P_1\quad ^3P_2$
2	422.642	236,607	$3s3p\ ^3P_2-3s3d\ ^3D_2$
10	422.511	236,680	$^3P_2\quad ^3D_3$
1	418.571	238,908	$^3P_1\quad ^3D_1$
7	418.443	238,981	$^3P_1\quad ^3D_2$
3	416.628	240,022	$^3P_0\quad ^3D_1$
15	230.745	433,379	$3s3d\ ^3D_3-3s4f\ ^3F_4$
12	230.706	433,452	$^3D_2\quad ^3F_3$
10	230.678	433,504	$^3D_1\quad ^3F_2$
15	199.921	500,198	$3s3p\ ^3P_2-3s4s\ ^3S_1$
10	198.975	502,576	$^3P_1\quad ^3S_1$
5	198.536	503,688	$^3P_0\quad ^3S_1$
1	156.566	638,708	$3s3p\ ^3P_2-3s4d\ ^3D_1$
3	156.543	638,802	$^3P_2\quad ^3D_2$
9	156.507	638,949	$^3P_2\quad ^3D_3$
4	155.985	641,087	$^3P_1\quad ^3D_1$
7	155.961	641,185	$^3P_1\quad ^3D_2$
5	155.715	642,199	$^3P_0\quad ^3D_1$

* Now at the University of Chicago.

[1] L. W. Phillips, Phys. Rev. **53**, 248 (1938).

[2] P. G. Kruger and L. W. Phillips, Phys. Rev. **55**, 352 (1939).

[3] R. F. Bacher and S. Goudsmit, *Atomic Energy States* (McGraw-Hill Book Co., 1932).

[4] I. S. Bowen, Phys. Rev. **39**, 8 (1932).

[5] I. S. Bowen and R. A. Millikan, Phys. Rev. **25**, 597 (1925).

[6] B. Edlén, Zeits. f. Physik **103**, 536 (1936).

TABLE III. *Classified lines of* Ca IX.

INT	λ(VAC.)	ν(CM⁻¹)	CLASSIFICATION
1	515.576	193,958	$3s3p\ ^3P_2-3p^2\ ^3P_1$
0*	512.060	195,290	$^3P_1\ \ ^3P_0$
0	507.105	197,198	$^3P_1\ \ ^3P_1$
4*	506.163	197,565	$^3P_2\ \ ^3P_2$
0*	503.261	198,704	$^3P_0\ \ ^3P_1$
1*	497.996	200,805	$^3P_1\ \ ^3P_2$
2Bl	378.550	264,166	$3s3p\ ^3P_2-3s3d\ ^3D_1$
2	378.382	264,283	$^3P_2\ \ ^3D_2$
20	378.089	264,488	$^3P_2\ \ ^3D_3$
15Bl	373.979	267,395	$^3P_1\ \ ^3D_1$
10	373.802	267,521	$^3P_1\ \ ^3D_2$
5	371.899	268,890	$^3P_0\ \ ^3D_1$
8	184.433	542,202	$3s3d\ ^3D_3-3s4f\ ^3F_4$
6	184.377	542,366	$^3D_2\ \ ^3F_3$
3	184.338	542,483	$^3D_1\ \ ^3F_2$
6	163.229	612,635	$3s3p\ ^3P_2-3s4s\ ^3S_1$
4	162.371	615,872	$^3P_1\ \ ^3S_1$
2	161.979	617,364	$^3P_0\ \ ^3S_1$
1	129.970	769,410	$3s3p\ ^3P_2-3s4d\ ^3D_2$
10	129.934	769,620	$^3P_2\ \ ^3D_3$
1	129.445	772,530	$^3P_1\ \ ^3D_1$
6	129.425	772,650	$^3P_1\ \ ^3D_2$
3	129.197	774,010	$^3P_0\ \ ^3D_1$
2	100.958	990,510	$3s3p\ ^3P_2-3s5d\ ^3D_3$
1	100.645	993,590	$^3P_1\ \ ^3D_2$

* Identified by Edlen.
Bl—Blend with another line.

tions with several other terms, it is now possible to make correct J value assignments in the pp' group. Since this group of lines is out of the range of the Illinois spectrograph, the intensities reported are those given by Bowen and Millikan.[5]

Whitford[7] has classified six lines taken from Ekefors'[8] K data as the pp' group in K VIII. The six lines he chose are the correct ones, but the J value assignments, and hence his term table, appear to be incorrect. The classification given in Table II is verified by other combinations with $3s3p\ ^3P$ term.

Four of the lines in the pp' group in Ca IX are given in Ekefors'[8] Ca data and were correctly identified by Edlén.[6] Two of the lines in the $3s3p\ ^3P-3s3d\ ^3D$ multiplet in Ca IX appear to be more intense than one would expect them to be. Inspection of the plate indicates that in reality these two lines are both blends with other calcium lines belonging to some other stage of ionization, and it has been so indicated in Table III.

[7] A. E. Whitford, Phys. Rev. **46**, 793 (1934).
[8] E. Ekefors, Zeits. f. Physik **71**, 53 (1931).

TABLE IV. *Classified lines of* Sc X.

INT	λ(VAC.)	ν(CM⁻¹)	CLASSIFICATION
1	343.100	291,460	$3s3p\ ^3P_2-3s3d\ ^3D_1$
1	342.877	291,650	$^3P_2\ \ ^3D_2$
10	342.509	291,963	$^3P_2\ \ ^3D_3$
1	338.115	295,757	$^3P_1\ \ ^3D_1$
5	337.888	295,956	$^3P_1\ \ ^3D_2$
2	335.815	297,783	$^3P_0\ \ ^3D_1$
10	151.001	662,247	$3s3d\ ^3D_3-3s4f\ ^3F_4$
5	150.941	662,511	$^3D_2\ \ ^3F_3$
3	150.902	662,682	$^3D_1\ \ ^3F_2$
8	135.921	735,721	$3s3p\ ^3P_2-3s4s\ ^3S_1$
4	135.131	740,020	$^3P_1\ \ ^3S_1$
2	134.766	742,025	$^3P_0\ \ ^3S_1$

TABLE V. *Term values.*

CONFIG.	Cl VI	K VIII	Ca IX	Sc X
$3s3p\ ^3P_0$	0	0	0	0
$^3P_1$	553	1,112	1,495	2,015
$^3P_2$	1,718	3,484	4,735	6,317
$3p^2\ ^3P_0$	136,713	176,701	196,785	
$^3P_1$	137,449	178,067	198,698	
$^3P_2$	138,650	180,640	202,300	
$3s3d\ ^3D_1$	181,698	240,036	268,890	297,777
$^3D_2$	181,713	240,092	269,017	297,969
$^3D_3$	181,741	240,164	269,223	298,280
$3s4s\ ^3S_1$	309,257	503,686	617,367	742,033
$3s4d\ ^3D_1$	411,721	642,197	774,017	
$^3D_2$	411,749	642,292	774,145	
$^3D_3$	411,800	642,433	774,355	
$3s4f\ ^3F_2$			811,368	960,459
$^3F_3$	431,742	673,543	811,388	960,480
$^3F_4$			811,420	960,527
$3s5d\ ^3D_2$	513,911		995,085	
$^3D_3$	513,942		995,245	

Term values for deep-lying states of the triplet system are given in Table V for all four ions. It has not been possible to make positive identification of any lines belonging to the singlet system, although the transition $3s^2\ ^1S-3s3p\ ^1P$ has been tentatively identified in Cl, K, and Ca. In Cl and K the separations of the $3s4f\ ^3F$ term calculated from observed lines turn out to be smaller than (or of the same order of magnitude as) the error of measurement. The ^3F level in these two ions has accordingly been reported as unresolved.

In conclusion, the authors wish to express their appreciation to Professor P. G. Kruger for his interest in the problem.

Magnetic Energy Constants of Dipolar Lattices

J. A. SAUER
*Department of Physics, Harvard University, Cambridge, Massachusetts**
(Received November 3, 1939)

In this paper, calculations are presented for the determination of the magnetic energy, assuming dipole-dipole interaction, of a macroscopic crystal possessing paramagnetic ions situated on a definite lattice structure. Parallel and antiparallel configurations are considered as well as specimens of different shapes and lattice structures of various types. It is shown that the configuration of lowest energy cannot be strictly defined but is dependent upon the shape of the specimen. This configuration is an antiparallel one for a sphere but changes to a parallel one for a long slender specimen. The transition between the two configurations appears to take place at a certain critical value of the demagnetization factor. For specimens above this value the theory accounts qualitatively for the low values of the magnetization found experimentally and for the apparent constancy of the Curie point with changes of shape. Below the critical value the theory implies a pure ferromagnetic state, which is not in agreement with experiment. Probable reasons for the discrepancy are given.

IN attempting to give a theoretical explanation of the anomalous experimental results of Kurti and Simon[1] on paramagnetic salts at low temperatures, the author found it desirable to first make a definite determination of the configuration of lowest energy. This paper is devoted to that end. The model used is a three-dimensional lattice of freely suspended magnetic dipoles, each of which is assumed to possess two orientations, either parallel or antiparallel to some given direction.

To compute the configurational energy, the exact method of attack would be to calculate by wave-mechanical means the energy levels of the assembly as a whole, considering the crystal as one large molecule. Owing, however, to the complexity of the wave equation when interaction between particles is considered, this method is prohibitive. Therefore we resort to a purely classical determination of the magnetic energy of the given three-dimensional configurations, assuming different possible arrangements of the paramagnetic ions. From spectroscopic analysis, as carried out by Beevers and Lipson,[2] the structure of such paramagnetic salts as iron alum is now well known. The unit cell is about 12 angstroms in length and the metal atoms are placed on a face-centered structure. It is due to these comparatively large distances apart that the exchange forces, which fall off exponentially with the distance, probably do not matter; whereas the magnetic forces, varying only as the inverse third power of the distance, become of prime importance.

The dependence of the energy upon the magnetic forces rather than the exchange forces between atoms makes the calculations much more complex as the magnetic interaction between two magnets, considering only dipole moments, depends on their orientation relative to the connecting line as well as on their distance apart. The classical expression for the interaction energy is

$$(\boldsymbol{\mu}_i \cdot \boldsymbol{\mu}_j) r_{ij}^{-3} - 3(\boldsymbol{\mu}_i \cdot \mathbf{r}_{ij})(\boldsymbol{\mu}_j \cdot \mathbf{r}_{ij}) r_{ij}^{-5}, \qquad (1)$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ are the strength of the magnetic dipoles and $\mathbf{r}_{ij}$ is the distance between the two. If the dipoles have the same direction as well as strength, the expression simplifies to $\mu^2 r^{-3}(1-3\cos^2\theta)$ where θ is the angle between the dipole direction and the line joining the two dipoles.

For a single row of dipoles the lowest energy is obtained when they are all parallel to one another and directed along the line joining them. For a double row or a two-dimensional set-up, this is no longer true. Here a direct evaluation of Eq. (1), summed over all dipoles, shows that the antiparallel alignment—one string pointing to the right, the next nearest string to the left, and so on gives the deepest energy.

For the three-dimensional solid, with magnetic dipoles situated at all lattice points and all

* The major part of this paper was written while the author was at Cambridge University, England.

[1] Kurti, Laine and Simon, Comptes rendus **204**, 675, 754 (1937).

[2] C. A. Beevers and H. Lipson, Proc. Roy. Soc. **148**, 664 (1935).

pointing in the same direction, the magnetic energy is given by

$$W=-\sum_i \sum_{l,m,n} \frac{\mu_i^2}{r^3} \frac{2l^2-m^2-n^2}{(l^2+m^2+n^2)^{5/2}}$$

$$=-\tfrac{1}{2}N^2\mu^2 \sum_{l,m,n} \frac{2l^2-m^2-n^2}{(l^2+m^2+n^2)^{5/2}}, \quad (2)$$

where N is the number of ions per cc and l, m, and n are integers running from $-\infty$ to ∞. This is an infinite summation of the type $1/r^3$ in three dimensions. It cannot be definitely evaluated as the series is logarithmically divergent. This is equivalent to saying that the value of the interaction energy will depend on the shape. For a spherical or cubic specimen, symmetry requirements show directly that W will be zero. For a specimen of any other shape it becomes necessary to resort to the Lorentz approximation. Lorentz considers a small sphere surrounding the dipole in question and assumes that the summation over all dipoles can be expressed as a summation within the sphere and an integration outside. The summation contributes nothing to the field at the center, provided the dipoles are arranged in symmetrical fashion. There will, however be a contribution for unsymmetrical arrangements. This point has been considered by Sauer and Temperly,[3] who show that the internal field may be taken as proportional to the degree of order of the dipole arrangement. The value of the integration over all the dipoles outside the small sphere is determined by considering the effect of the surface charge produced by these outside magnets. On the surface of the small sphere itself there exists a charge density which contributes to the field at the center a factor of $(4\pi/3)I$, where I is the intensity of magnetization per unit volume; the additional field due to the surface charge on the external boundary of the specimen is brought in by what is commonly referred to as the demagnetization factor. For a sphere the demagnetization factor of $-(4\pi/3)$ exactly annuls the local field factor of $+(4\pi/3)$ so that the energy is given correctly by the Lorentz approximation.

For parallel alignment of the elementary dipoles the magnetic interaction energy, using the Lorentz approximation, can be obtained for any external shape of specimen provided the appropriate demagnetization factor is given or can be determined. But for other alignment configurations such as the antiparallel one, the Lorentz method fails and it becomes necessary, in order to get a reasonable approximation to the correct magnetic energy, to take a direct sum result over a large number of dipoles. All direct sum results hereinafter given will be determined by taking into account the effect of all dipoles within a small cube surrounding the dipole in question and containing approximately 125 atoms. Careful attention must be paid to algebraic signs in carrying out this summation as antiparallel dipoles give an interaction energy of opposite sign to that of parallel dipoles. For purposes of direct summation the equation for W can conveniently be written as

$$W=-\tfrac{1}{2}N^2\mu^2 \sum_\gamma \sum_l (2l^2-\gamma^2)/(l^2+\gamma^2)^{5/2}. \quad (3)$$

Here γ is the right-angle distance from the center dipole to whichever string is being considered. Thus if we were determining the energy for an antiparallel arrangement in a simple cubic crystal, γ would take the values 1, $\sqrt{5}$, $\cdots$ for the antiparallel dipoles, $\sqrt{2}$, 2, $\sqrt{8}$, $\cdots$ for the parallel dipoles, while W has the sign given by Eq. (3) for the parallel rows and changes sign for the antiparallel.

Although the paramagnetic ions are located on a face-centered lattice, it appears worth while for reference purposes to include the energy calculations for the simple cubic structure and the body-centered structure as well. In the case of the body- and face-centered structures, it is possible that a deeper energy might be obtained by assuming the magnets directed along the body or face diagonal rather than along the axis of the unit cell. Such indeed proves to be the case. We also distinguish between two different possible types of antiparallel alignment which we shall refer to as "A" and "B." In arrangement "A" the rows closest to one another are alternately directed. For example, in the body-centered cubic lattice the corner dipoles would all point forward and the body-centered dipoles would point backward. In arrangement "B," the nearest dipoles situated on a plane perpendicular to the

[3] J. A. Sauer and H. N. V. Temperly, Proc. Roy. Soc. (in print).

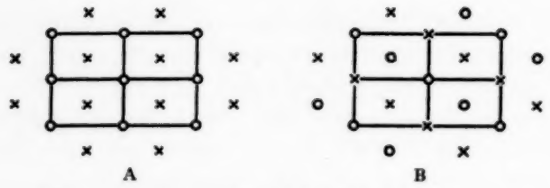

FIG. 1. Showing sections taken at right angles to direction of dipoles.

axis are alternately directed; in the example given above, this means that the corner dipoles are themselves arranged in antiparallel fashion and similarly for the body-centered ones. Any possible ambiguity arising from these definitions will be compensated by illustrating, wherever necessary, the difference between the two arrangements. The magnetic interaction energy for each of three alignments, parallel, antiparallel "A," and antiparallel "B" will now be obtained for different shaped specimens.

INTERACTION ENERGY FOR A SPHERICAL SPECIMEN

(*1*) *Parallel Alignment:* The magnetic energy for a simple cubic lattice structure is found to be zero whether the determination is made directly from Eq. (2) or by means of the Lorentz method. This result still holds good for the body-centered and face-centered lattices, as they consist, respectively, of two and four interlocked cubic lattices.

(*2*) *Antiparallel "A" and "B":* It is necessary here to consider each one of the three lattice structures separately.

Simple Cubic Structure: Evaluation of Eq. (3), taking into account the interaction of the nearest 125 dipoles, leads to an energy value of $W=-2.7N^2\mu^2$.

Body-centered Structure: We consider two different possibilities of dipole direction, along edge of unit cell or along body diagonal. For each of these chosen directions, the magnetic energy is calculated for both types of antiparallel alignment. The two configurations "A" and "B" are shown in Fig. 1, the circles indicating parallel dipoles and the crosses antiparallel ones. The interaction energy for the (001) direction and with arrangement "A" is zero, while with arrangement "B," the result is -1.35. For the diagonal direction, W is -1.75 for "A" and "B."

Face-centered Structure: Fig. 1 may again be used for noting the essential difference between the two types of alternating alignment. The approximate values of the energy are 2.2 for arrangement "A" and dipole direction along cube edge, and -1.1 for the same direction with "B." With face-diagonal direction of the magnets, W decreases to -1.1 for "A" and -1.8 for "B."

INTERACTION ENERGY FOR AN ELLIPSOIDAL SPECIMEN

(*1*) *Parallel Alignment:* For this case the energy may be expressed as

$$W=-0.5\sum\mu[(4\pi/3)-(\text{D.F.})]I \qquad (4)$$

where by D.F. is meant the demagnetization factor. The values obtained will differ from the energies in the case of the sphere by having the proper demagnetizing factor, depending on the axis ratio of the ellipsoid, inserted in place of the $4\pi/3$ of the sphere. For example, a specimen with an axis ratio of $r=3$ possesses a demagnetization factor of 1.37 and gives rise to an interaction energy of 1.9. As the axis ratio approaches infinity, D.F. approaches zero and W is then given by $0.5\sum\mu\cdot(4\pi/3)N\mu=2.1N^2\mu^2=2.1$ in units of $N^2\mu^2$. The results are the same for all three types of crystal lattice structure.

(*2*) *Alternating Arrangements:* The energy values as given above for a spherical specimen are true here as well. This is so because in the alternating arrangement of dipoles there exists no charge density on the surface and hence the shape of the boundary does not matter.

For simplicity of reference and comparison, the foregoing results, together with some other energy values as yet unmentioned, have been assembled in Table I. Some of these magnetic energy values have already been given by Van Vleck,[4] but as no details of the calculation and also no explicit values for the possible diagonal directions appeared in his paper, it was thought desirable to collect all the values in one Table. In addition to the energies for the parallel configuration, the antiparallel "A," and the antiparallel "B," Table I includes the magnetic energy for a rather special configuration, pictured in Fig. 2. This arrangement of dipoles might be

[4] J. H. Van Vleck, J. Chem. Phys. 5, 556 (1937).

termed a nonmagnetic, disordered state of the crystal as compared to the parallel arrangement, which is a magnetic, ordered state and the antiparallel arrangement, which is also ordered but nonmagnetic. This particular disordered state is only one of a number of possible disordered arrays, yet it is of interest to compare its energy values with those of the ordered arrangements.

Examination of Table I allows us to draw the following conclusions:

(1) The parallel alignment gives an energy independent of the type of lattice structure but dependent upon the shape of the specimen. That is, for a given shape, the energy is the same whether the dipoles are arranged on a body-centered, face-centered, or simple cubic lattice yet a change in specimen shape from a sphere to a long slender specimen changes the energy from the value zero to -2.1.

(2) Conversely, the antiparallel configurational energy is dependent upon the type of lattice structure but independent of the shape. For example, the energy changes in value from -2.7 to 1.1 in going from a simple cubic lattice to a face-centered one, while a change in the external boundary leaves the values unaltered. Comparing the two types of antiparallel alignment, we see that the lower energy is in general associated with type "B" rather than "A."

(3) Dipoles set along the direction of the face diagonal for the face-centered lattice and along the direction of the body-diagonal for the body-centered lattice, lead to deeper energy values than when directed along the cube axis. This is what we should expect for changing direction from cube axis to body or face diagonal reduces the distance between adjacent dipoles of a parallel string and hence causes an increase in the magnetic energy.

(4) For parallel alignment dipoles tend to set themselves in such a direction that the demagnetizing factor is a minimum. In other words, in an ellipsoidal specimen, the dipoles would prefer to align along the long axis rather than the short.

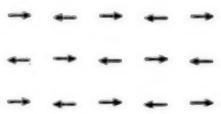

FIG. 2.

TABLE I. *Energy constants. All direct sum values (antiparallel ones) given to within 0.05. a—simple cubic, dipole direction along (001); b—body-centered, dipole direction along (001); c—body-centered, dipole direction along (111); d—face-centered, dipole direction along (001); e—face-centered, dipole direction along (011).*

DIPOLE ARRANGEMENTS	LATTICE STRUCTURE	AXIS RATIOS	ENERGY CONSTANT
Parallel	*a, b, c, d, e*	$\frac{1}{3}$	1.91
		1	0.0
		3	−1.41
		4	−1.62
		5	−1.75
		6	−1.83
		∞	−2.1
Antiparallel "A"	*a*	any	−2.7
	b		0.0
	c		−1.75
	d		2.2
	e		−1.1
Antiparallel "B"	*a*	any	−2.7
	b		−1.35
	c		−1.75
	d		−1.1
	e		−1.8
Disordered	*a*		0.0

(5) The actual configuration of minimum energy is not strictly defined but is dependent upon the shape. For a sphere or an ellipse with low axis ratio, the antiparallel arrangement is definitely the lower. For a long slender specimen or an ellipse with high axis ratio ($r>6$ for the face-centered lattice) the parallel configuration becomes the more stable.

(6) A comparison of the values from Table I for the parallel and the disordered configurations indicates that a spherical specimen, in the absence of any external field, has equal energies in the completely ordered state and in the completely disordered state. This, perhaps, may not be an expected relation; yet when it is recalled that for a spherical specimen the Lorentz and the demagnetizing terms cancel, thus causing the magnetic interaction energy to vanish, the correlation appears more natural. The recorded energy value for the disordered array was calculated only for a simple cubic lattice but it appears likely that the same result would be obtained for the other lattice types.

The three-dimensional model studied in this paper appears to eliminate the chance of a spherical specimen becoming spontaneously magnetized as a result of pure dipole-dipole coupling. The ferromagnetic state would be possible,

though, for a long slender specimen. Now it is well known that the usual Ising model, to which our model roughly corresponds, will never give ferromagnetism. This result, however, holds only in the case of a linear chain. Peierls[5] has generalized Ising's model to include the two- and three-dimensional cases and he finds that ferromagnetism is possible. It is not correct though to assume, in our model, that even a long slender specimen will necessarily be ferromagnetic. For even if the lowest energy level of the crystal does correspond to the parallel state, there may exist other nonparallel states of only slightly higher energy values; and the greater number of complexions of these nonparallel states might easily be sufficient to overcome the energy advantage of the parallel level.

By taking the antiparallel configuration as the stable one at low temperatures, and by using the energy values given in Table I, it is possible to determine the Curie points. This is done in a paper by Sauer and Temperly[3] who show that the Curie point, for all specimens in which the antiparallel state is the stable one, i.e., for ellipsoidal shapes with r less than 6, the critical temperature is independent of the shape. This appears to be in confirmation with the experimental results as Dr. Simon has reported that for a sphere as well as long samples, there is very little change in the Curie point.

In view of our results, the Curie point for a sphere is now to be interpreted as the temperature at which a transition takes place between a disordered state and the ordered antiparallel state. Heretofore, it has usually been taken as an indication of a conversion from the disordered state to the ordered parallel one; because of this interpretation, one expected the measured magnetization to be in the neighborhood of its saturation value. All experiments, however, show it to be very much smaller. This has always been considered one of the very puzzling features of the experimental results but the mystery is readily removed when we recognize that the ordered state below the Curie point is not magnetic but essentially nonmagnetic.

For an ellipsoidal specimen with $r>6$ the reason for a low value of the magnetization is not so apparent, as here the Curie point should correspond as usual to the transition to the parallel saturated state, provided, of course, that our simple treatment is adequate. Before definite conclusions can be drawn there is a great need for more experimental evidence, especially in the region of $r>6$. The low value of the magnetization is inferred, at present but from a single experimental value on iron alum with an r ratio of 6. Still there does appear to be a definite discrepancy between theory and experiment. The probable reason for this discrepancy is the too definite postulation of the existence of the parallel state as the stable one on the simple evidence that its energy value is slightly less than that of the antiparallel state. In a rigorous treatment the effect of the greater statistical weight of the nonparallel states must needs be considered.

It requires very little change to convert the energy constants of Table I to conform to the case of electrical dipole-dipole interaction instead of magnetic. Nothing in the electrical case corresponds to the demagnetization factor in the magnetic, hence the D.F. term of Eq. (4) must vanish. This converts all the different energy constants for parallel alignment to the single value -2.1. The antiparallel configurational energies remain the same. A comparison now leads to the conclusion that, for the simple cubic lattice, the antiparallel arrangement still gives the deeper energy. For the face-centered and body-centered lattices, however, the electric dipole coupling leads to different results than the magnetic. It indicates that, of the various arrangements considered, the parallel state is always the most stable regardless of the shape of the specimen. This does not imply electric saturation for the same reasons as given above for the magnetic case.

I wish to express my gratitude to Professor Fowler for suggesting this problem and for his generous assistance. I am also indebted to Mr. A. H. Wilson and Professor Van Vleck for helpful discussions.

[5] R. Peierls, Proc. Camb. Phil. Soc. 32, 471 (1936).

General Relativity and Flat Space. I

N. ROSEN
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received October 30, 1939)

Within the framework of the general theory of relativity, it is proposed to introduce at each point of space-time a Euclidean metric tensor $\gamma_{\mu\nu}$ in addition to the usual Riemannian metric tensor $g_{\mu\nu}$. In this way one imparts tensor character to quantities which in the usual form of the theory do not have it. For example, one can obtain a gravitational energy-momentum density tensor in place of the usual pseudo-tensor. Furthermore one can impose four additional covariant conditions on the gravitational field and thus restrict the form of the solution for the field corresponding to a given physical situation.

§1

IN the general theory of relativity, Einstein[1] makes use of Riemannian geometry, characterized by the existence of an invariant interval or line element ds, given by

$$ds^2=g_{\mu\nu}dx^\mu dx^\nu, \tag{1}$$

where $g_{\mu\nu}$ is the metric tensor and (x^1,x^2,x^3,x^4) are the coordinates in an arbitrary coordinate system. The metric tensor is determined by the law of gravitation:

$$G_{\mu\nu}=-8\pi T_{\mu\nu}, \tag{2}$$

where $T_{\mu\nu}$ is the energy-momentum density tensor, and

$$G_{\mu\nu}\equiv R_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}R, \tag{3}$$

with $R_{\mu\nu}$ the contracted Riemann-Christoffel tensor.

The tensor $G_{\mu\nu}$ satisfies the identity

$$G_\mu{}^\alpha{}_{;\alpha}\equiv 0, \tag{4}$$

where (;) denotes covariant differentiation based on $g_{\mu\nu}$ (g-differentiation). This permits setting $G_{\mu\nu}$ proportional to $T_{\mu\nu}$ which is taken to satisfy the equation (expressing the laws of motion of matter or energy)

$$T_\mu{}^\alpha{}_{;\alpha}=0. \tag{5}$$

§2

It is now proposed to introduce at every point of space-time, along with the metric tensor $g_{\mu\nu}$, a second metric tensor $\gamma_{\mu\nu}$ corresponding to flat space, i.e., for which the Riemann-Christoffel tensor vanishes identically everywhere. This may be interpreted in various ways. For instance, we may suppose that we map the Riemannian space with metric $g_{\mu\nu}$ on a flat space with metric $\gamma_{\mu\nu}$ and assign the same values to the coordinates of corresponding points. Or we may say that the two metrics side by side represent a comparison of the given space with the space one would have if the gravitational field were removed. It should be emphasized that in introducing the $\gamma_{\mu\nu}$, we are not postulating here any new properties of the space.

In this way one can define a Euclidean line element analogous to (1)

$$d\sigma^2=\gamma_{\mu\nu}dx^\mu dx^\nu. \tag{6}$$

Moreover one can now define covariant differentiation based on $\gamma_{\mu\nu}$ (γ-differentiation), which will be denoted by (,). Since the Riemann-Christoffel tensor formed from $\gamma_{\mu\nu}$ has been assumed to vanish, it follows that one can interchange the order of γ-differentiation, so that the latter obeys all the rules of ordinary differentiation, except that the γ-derivatives of $\gamma_{\mu\nu}$ vanish. Indeed it is always possible to choose a coordinate system (special relativity) such that the components $\gamma_{\mu\nu}$ are constant and γ-differentiation reduces to ordinary; but it may not always be convenient to choose such a system.

Let us next consider $\Delta^\lambda{}_{\mu\nu}$ defined by

$$\left\{\begin{matrix}\lambda\\ \mu\ \ \nu\end{matrix}\right\}=\Gamma^\lambda{}_{\mu\nu}+\Delta^\lambda{}_{\mu\nu}, \tag{7}$$

where $\{^\lambda_{\mu\nu}\}$ is the Christoffel 3-index symbol used in g-differentiation and $\Gamma^\lambda{}_{\mu\nu}$ that occurring in γ-differentiation. This can be shown to be a tensor and is found to be given by

$$\Delta^\lambda{}_{\mu\nu}=\tfrac{1}{2}g^{\lambda\alpha}(g_{\mu\alpha,\nu}+g_{\nu\alpha,\mu}-g_{\mu\nu,\alpha}). \tag{8}$$

[1] A. Einstein, Ann. d. Physik **49**, 769 (1916).

If in the usual expression[2] defining $R_{\mu\nu}$, one substitutes (7), it is found that, as a consequence of the Euclidean character of $\gamma_{\mu\nu}$, one can write

$$R_{\mu\nu}=-\Delta^{\alpha}{}_{\mu\nu,\alpha}+\Delta^{\alpha}{}_{\alpha\mu,\nu}-\Delta^{\alpha}{}_{\alpha\beta}\Delta^{\beta}{}_{\mu\nu}+\Delta^{\alpha}{}_{\beta\mu}\Delta^{\beta}{}_{\alpha\nu}. \quad (9)$$

This relation is interesting in that $R_{\mu\nu}$ is seen to be obtainable from the tensor $\Delta^{\lambda}{}_{\mu\nu}$ by tensor operations, i.e., $R_{\mu\nu}$ is a tensor function of $g_{\mu\nu}$.

§3

Comparing (9) with the usual expression for $R_{\mu\nu}$, one sees that $\{{}^{\lambda}_{\mu\nu}\}$ has been replaced by $\Delta^{\lambda}{}_{\mu\nu}$ and ordinary differentiation by γ-differentiation. This type of correspondence turns out to be quite general. It appears that one can rewrite all the quantities occurring in relativity theory so that $\{{}^{\lambda}_{\mu\nu}\}$ is replaced by $\Delta^{\lambda}{}_{\mu\nu}$, an ordinary derivative by a γ-derivative (in particular, $\partial g_{\mu\nu}/\partial x^{\sigma}$ by $g_{\mu\nu,\sigma}$) and $(-g)^{\frac{1}{2}}$ by

$$\kappa=(g/\gamma)^{\frac{1}{2}}, \quad (10)$$

where g and γ are the determinants of $g_{\mu\nu}$ and $\gamma_{\mu\nu}$, respectively. Since $g_{\mu\nu,\sigma}$ is a tensor and κ a scalar, it follows that some quantities, on being rewritten in this way, take on a tensor character, although they do not have it in the usual form of the theory.

To illustrate this point, the following example may be useful to compare with the analogous calculation usually made.[3] One can show that the gravitational Eqs. (2) are connected with a variational principle

$$\delta\int L(-g)^{\frac{1}{2}}d\tau\left(\equiv\delta\int L\kappa(-\gamma)^{\frac{1}{2}}d\tau\right)$$

$$=\int G^{\mu\nu}\delta g_{\mu\nu}(-g)^{\frac{1}{2}}d\tau, \quad (11)$$

where

$$d\tau=dx^1dx^2dx^3dx^4,$$

$$L=g^{\mu\nu}(\Delta^{\alpha}{}_{\beta\mu}\Delta^{\beta}{}_{\alpha\nu}-\Delta^{\alpha}{}_{\alpha\beta}\Delta^{\beta}{}_{\mu\nu}), \quad (12)$$

and only the components of $g_{\mu\nu}$ are varied (with the variations vanishing on the boundaries of the region of integration). This function L is a scalar. From it one can derive a gravitational energy-momentum density tensor (analogous to the pseudo-tensor of Einstein). One writes (Eq. (4))

$$G_{\mu}{}^{\alpha}{}_{,\alpha}+G^{\beta}{}_{\mu}\Delta^{\alpha}{}_{\beta\alpha}-G^{\alpha}{}_{\beta}\Delta^{\beta}{}_{\mu\alpha}=0, \quad (13)$$

or, since

$$\Delta^{\alpha}{}_{\alpha\beta}=\kappa_{,\beta}/\kappa, \quad (14)$$

$$(\kappa G^{\alpha}{}_{\mu})_{,\alpha}-\tfrac{1}{2}\kappa G^{\alpha\beta}g_{\alpha\beta,\mu}=0. \quad (15)$$

Throughout this paper, unless otherwise specified, indices are raised and lowered by means of $g_{\mu\nu}$.

Now (11) can be shown to be equivalent to the relation

$$\kappa G^{\mu\nu}=\partial(L\kappa)/\partial g_{\mu\nu}-[\partial(L\kappa)/\partial g_{\mu\nu,\sigma}]_{,\sigma}, \quad (16)$$

and from this one obtains (using integration by parts)

$$\kappa G^{\alpha\beta}g_{\alpha\beta,\mu}=\{L\kappa\delta^{\lambda}{}_{\mu}-[\partial(L\kappa)/\partial g_{\alpha\beta,\lambda}]g_{\alpha\beta,\mu}\}_{,\lambda}, \quad (17)$$

which is to be substituted in (15). If we now define as the gravitational energy-momentum density tensor,

$$t_{\mu}{}^{\lambda}=\frac{1}{16\pi}\{L\delta_{\mu}{}^{\lambda}-[\partial(L\kappa)/\kappa\partial g_{\alpha\beta,\lambda}]g_{\alpha\beta,\mu}\}, \quad (18)$$

we can write (15) or (4) as

$$[\kappa(T_{\mu}{}^{\lambda}+t_{\mu}{}^{\lambda})]_{,\lambda}=0. \quad (19)$$

Or, taking

$$T'_{\mu}{}^{\lambda}=\kappa T_{\mu}{}^{\lambda}, \quad t'_{\mu}{}^{\lambda}=\kappa t_{\mu}{}^{\lambda}, \quad (20)$$

we can write

$$(T'_{\mu}{}^{\lambda}+t'_{\mu}{}^{\lambda})_{,\lambda}=0. \quad (21)$$

By a suitable coordinate transformation, one can put the left-hand member into the form of an ordinary divergence. The fact that $t'_{\mu}{}^{\lambda}$ is a tensor serves to remove some objections that had been raised in the past.[4]

§4

It is to be seen that, in general, the quantities in the usual form of the relativity theory can be considered as arising from the corresponding ones here, through the special choice of $\gamma_{\mu\nu}$,

$$\gamma_{\mu\nu}=\text{constant}. \quad (22)$$

[2] E.g., L. P. Eisenhart, *Riemannian Geometry* (Princeton, 1926), p. 21.

[3] Reference 1, p. 804, and H. Weyl, *Raum Zeit Materie* (Berlin, 1923), p. 272.

[4] H. Bauer, Physik. Zeits. **19**, 163 (1918).

However, if (22) is to hold for all coordinate systems, $\gamma_{\mu\nu}$ cannot be a tensor. Consequently the quantities which depend on $\gamma_{\mu\nu}$ lose their tensor character. Thus the gravitational tensor t_μ^λ of (18) goes over into the pseudo-tensor in the Einstein theory. On the other hand, $R_{\mu\nu}$ is independent of $\gamma_{\mu\nu}$ and hence remains a tensor. Obviously, if one wishes to maintain the tensor character of quantities depending on $\gamma_{\mu\nu}$ one must give up (22) and allow $\gamma_{\mu\nu}$ to transform as a tensor under coordinate transformations.

§5

Although $g_{\mu\nu}$ and $\gamma_{\mu\nu}$ have been considered as existing side by side, so far nothing has been said about any relation between them. Actually one would expect to have some relation between them, for it seems reasonable that, if the gravitational field is made to vanish, $g_{\mu\nu}$ should go over into $\gamma_{\mu\nu}$. Such a relation can be obtained by imposing four additional covariant conditions on the field. This can be done because of the set of identities (4) existing among the field equations. In working with the linear approximation of the gravitational equations, Einstein[5] similarly added four conditions on the field for the purpose of eliminating (or reducing) apparent fields arising from infinitesimal coordinate transformations. It has hitherto not been possible to set up corresponding (covariant) conditions for the exact equations. In the present form of the theory this can be done, and it is these equations that provide the relation between $g_{\mu\nu}$ and $\gamma_{\mu\nu}$. They should serve to remove, or at least restrict, the ambiguity in the form of a solution arising from the possibility of coordinate transformations which go over into the identical transformation in the absence of a field.

For example, the static spherically symmetric solution of the gravitational equations in free space can be written in a number of forms depending on the choice of the radial variable r, it being asserted that one cannot know which choice corresponds to the variable r in flat space. The additional conditions would serve to single out one of the various possibilities.

One reasonable set of conditions can be obtained from the following considerations:

One can rewrite the expression (9) for $R_{\mu\nu}$ in the form

$$R_{\mu\nu}=\tfrac{1}{2}g^{\alpha\beta}g_{\mu\nu,\alpha\beta}-\tfrac{1}{2}S_{\mu;\nu}-\tfrac{1}{2}S_{\nu;\mu}-\Delta^{\alpha}{}_{\mu\beta}\Delta^{\alpha^*}{}_{\nu\beta^*}-\Delta^{\alpha}{}_{\mu\beta}\Delta^{\nu^*}{}_{\alpha\beta^*}-\Delta^{\alpha}{}_{\nu\beta}\Delta^{\mu^*}{}_{\alpha\beta^*}, \quad (23)$$

where

$$S_\mu\equiv\Delta^{\alpha}{}_{\mu\alpha^*}\equiv g^{\alpha\beta}g_{\alpha\mu,\beta}-\kappa_{,\mu}/\kappa, \quad (24)$$

and an asterisk with an index indicates raising or lowering it with $g_{\mu\nu}$.

This suggests taking as additional conditions

$$S_\mu=0. \quad (25)$$

In consequence, the right-hand member of (23) is simplified, only the first term now containing second derivatives. It can be verified that the conditions (25) in first approximation are the same as those imposed by Einstein in the case of the linear equations.

The static spherically symmetric solution of the field equations in free space

$$R_{\mu\nu}=0 \quad (26)$$

together with (25), is found to be

$$ds^2=-\frac{r+m}{r-m}dr^2-(r+m)^2(d\theta^2+\sin^2\theta d\phi^2)+\frac{r-m}{r+m}dt^2, \quad (27)$$

with the usual notation, if one takes

$$d\sigma^2=-dr^2-r^2(d\theta^2+\sin^2\theta d\phi^2)+dt^2. \quad (28)$$

Another possible set of conditions can be obtained based upon taking

$$\kappa=1. \quad (29)$$

This serves to simplify many of the formulas. It should be noted that this differs very essentially from the condition $g=-1$, often used, in that it is not affected by coordinate transformations. Three further conditions are required. One might take (as being equivalent to three conditions)

$$S_{\mu,\nu}-S_{\nu,\mu}=0. \quad (30)$$

In this case if one takes for $d\sigma^2$ the expression (28), the static spherically symmetric solution has the Schwarzschild[6] form:

$$ds^2=-(1-2m/r)^{-1}dr^2-r^2(d\theta^2+\sin^2\theta d\phi^2)+(1-2m/r)dt^2. \quad (31)$$

However this solution leads to the difficulty pointed out by Schrödinger[7] (a not very serious difficulty, however[8]) that

[5] A. Einstein, Berl. Ber., p. 688 (1916).

[6] K. Schwarzschild, Berl. Ber., p. 189 (1916).

[7] E. Schrödinger, Physik. Zeits. 19, 4 (1918).

[8] A. Einstein, Physik. Zeits. 19, 115 (1918); Berl. Ber., p. 448 (1918).

the components of t_μ^ν all vanish for $r>2m$. If one wishes to avoid this, one must give up (29) and (30).

§6

Enough has been given to show that there are advantages from the formal point of view in introducing the Euclidean $\gamma_{\mu\nu}$ into the general relativity theory. It imparts tensor character to quantities which otherwise do not have it, and allows additional conditions to be imposed on the field so as to restrict the form of the solution for a given physical situation.

In conclusion, it is necessary to point out that, having once introduced $\gamma_{\mu\nu}$ into the theory, one has a great number of new tensors and scalars at one's disposal. One can, therefore, set up other field equations than (2). It is possible that some of these may be more satisfactory for the description of nature. Further investigation is here required.

JANUARY 15, 1940 PHYSICAL REVIEW VOLUME 57

General Relativity and Flat Space. II

N. ROSEN
Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received October 30, 1939)

The possibility is considered of interpreting the formalism of the general theory of relativity in terms of flat space, the fundamental tensor $g_{\mu\nu}$ being regarded as describing the gravitational field but having no direct connection with geometry. The resulting theory in general leads to the same predictions as the Einstein theory, but there are cases where the predictions differ. The present theory may explain the principal results obtained by D. C. Miller in his "ether-drift" experiments. The implications of the theory for cosmology are briefly touched upon.

§1

IN a previous paper[1] (hereafter referred to as I) it was shown that it is useful to introduce into the general theory of relativity the concept of the existence at each point of space-time of a Euclidean metric tensor $\gamma_{\mu\nu}$ in addition to the usual Riemannian metric tensor $g_{\mu\nu}$. From the standpoint of the general theory of relativity, one must look upon $\gamma_{\mu\nu}$ as a fiction introduced for mathematical convenience. However, the question arises whether it may not be possible to adopt a different point of view, one in which the metric tensor $\gamma_{\mu\nu}$ is given a real physical significance as describing the geometrical properties of space, which is therefore taken to be flat, whereas the tensor $g_{\mu\nu}$ is to be regarded as describing the gravitational field.[2]

It has been pointed out in I that the introduction of $\gamma_{\mu\nu}$ leads to the possibility of other laws than those adopted in general relativity. In the present paper, however, no attempt will be made to change the laws of the latter, since they form a self-consistent system and have proved to be quite satisfactory for the description of large scale phenomena, at any rate.

§2

As far as the field equations are concerned, it is immaterial what interpretation one gives to the variables involved. This is not the case with the equations of motion. Let us therefore consider the law of motion for a particle in the field. The latter is given in the general theory of relativity by the equation of the geodesic[3]

$$\frac{d^2x^\mu}{ds^2}+\left\{\begin{matrix}\mu\\ \alpha\ \beta\end{matrix}\right\}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds}=0, \qquad (1)$$

where ds is the line element, defined in terms of the tensor $g_{\mu\nu}$, (I (1)). Let us now introduce as independent variable the Euclidean line element

[1] N. Rosen, Phys. Rev. **57**, 146 (1940).

[2] In some respects this resembles the theory of gravitation proposed by Nordström (cf. report by M. v. Laue, Jahrbuch f. Rad. u. El. **14**, 263 (1917)). It will be seen that there are important differences, however.

[3] A. Einstein, Ann. d. Physik **49**, 769 (1916).

$d\sigma$, defined in terms of the tensor $\gamma_{\mu\nu}$, (I (6)), instead of ds. After canceling out a factor $d\sigma/ds$, the equation can be written

$$\frac{d}{d\sigma}\left(\frac{d\sigma}{ds}\frac{dx^\mu}{d\sigma}\right)+\left\{\begin{matrix}\mu\\ \alpha\ \beta\end{matrix}\right\}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}\frac{d\sigma}{ds}=0. \tag{2}$$

This can be rewritten

$$\frac{d^2x^\mu}{d\sigma}+\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}=\lambda\frac{dx^\mu}{d\sigma}-\Delta^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}, \tag{3}$$

where $$\lambda\equiv\frac{d^2s}{d\sigma^2}\Big/\frac{ds}{d\sigma}\equiv-\frac{d^2\sigma}{ds^2}\Big/\left(\frac{d\sigma}{ds}\right)^2, \tag{4}$$

and $\Gamma^\mu_{\alpha\beta}$ and $\Delta^\mu_{\alpha\beta}$ are given by I, (7) and (8).

If we multiply (3) by $\gamma_{\mu\nu}dx^\nu/d\sigma$, the left-hand member vanishes as a consequence of the definition of $d\sigma$. Hence it is necessary that

$$\lambda=\gamma_{\tau\nu}\Delta^\tau_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}\frac{dx^\nu}{d\sigma}, \tag{5}$$

which gives the relation between s and σ. If this expression for λ is substituted into (3), one gets an equation from which s has been completely eliminated.

We can get the equation in a more familiar form if we go back to (2), multiply it through by a constant m_0, and write

$$m\equiv m_0 d\sigma/ds, \tag{6}$$

where m will be considered as the "proper mass" of the particle. Then the equation can be written as

$$\frac{d}{d\sigma}\left(m\frac{dx^\mu}{d\sigma}\right)+m\Gamma^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}=-m\Delta^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}. \tag{7}$$

On the left we have the absolute derivative with respect to σ of

$$p^\mu\equiv m\frac{dx^\mu}{d\sigma}\left(\equiv m_0\frac{dx^\mu}{ds}\right), \tag{8}$$

which is the energy-momentum vector. Hence the right-hand member is to be interpreted as the gravitational force. We see from (7) that, although inertial and gravitational mass are equal, they are not necessarily constant.

If we define the scalar

$$w^2=g_{\mu\nu}\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}, \tag{9}$$

then (6) can be written

$$m=m_0/w \tag{10}$$

as the law of dependence of proper mass on the gravitational field.

Conversely, if we start from (7), multiply by $g_{\mu\nu}dx^\nu/d\sigma$ and make use of I (7), we get

$$\tfrac{1}{2}mdw^2/d\sigma=-w^2dm/d\sigma, \tag{11}$$

which is equivalent to (10).

Thus we see that, from the standpoint of flat space, one can put the equations of motion of a particle into the Newtonian form (generalized to four dimensions) if we assign to the particle a variable mass. In a coordinate system in which the particle is at rest the dependence of mass on the field is given by

$$m=m_0(\gamma_{44}/g_{44})^{\frac{1}{2}}. \tag{12}$$

In a weak gravitational field this can be written in first approximation[4]

$$m=m_0(1-\varphi), \tag{13}$$

where φ is the gravitational potential.

In the case of a light ray we have the additional condition (in the presence of a gravitational field)

$$w=0. \tag{14}$$

We therefore let m_0 vanish also, but take m, as given by (10), to be finite and constant. This allows one to cancel out the m's in (7). That (14) is consistent with (7) is to be seen from (11).

§3

To see the connection between Eq. (7) and the gravitational field equations, (I (3)), let us suppose we have a stream of particles each of which moves according to (7). Let n_0 be the proper particle density and let us assume that all quantities that we may need are smooth functions of the coordinates. We write (7) in the form

$$\left(m\frac{dx^\mu}{d\sigma}\right)_{,\alpha}\frac{dx^\alpha}{d\sigma}=-m\Delta^\mu_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma} \tag{15}$$

and multiply through by n_0. On the left we have

$$\left(mn_0\frac{dx^\mu}{d\sigma}\frac{dx^\alpha}{d\sigma}\right)_{,\alpha}-m\frac{dx^\mu}{d\sigma}\left(n_0\frac{dx^\alpha}{d\sigma}\right)_{,\alpha}.$$

[4] Since it is shown, reference 3, that for a weak field $g_{44}\doteq 1+2\varphi$ in Galilean coordinates.

The scalar $(n_0(dx^\alpha/d\sigma))_{,\alpha}$ can be evaluated in a Galilean coordinate system. Thus

$$\left(n_0\frac{dx^\alpha}{d\sigma}\right)_{,\alpha}=\left(n_0\frac{dt}{d\sigma}\frac{dx^\alpha}{dt}\right)_{,\alpha}=\left(n\frac{dx^\alpha}{dt}\right)_{,\alpha}, \quad (16)$$

where t is the time coordinate and

$$n\equiv n_0 dt/d\sigma=n_0/(1-v^2)^{\frac{1}{2}} \quad (17)$$

is the particle density in a reference frame with respect to which the particles at the point in question have a velocity v, the velocity of light being taken as unity. Hence

$$\left(n_0\frac{dx^\alpha}{d\sigma}\right)_{,\alpha}=\frac{\partial n}{\partial t}+\frac{\partial(nv_x)}{\partial x}+\frac{\partial(nv_y)}{\partial y}+\frac{\partial(nv_z)}{\partial z} \quad (18)$$

and this vanishes (equation of continuity) if the number of particles is conserved, which we shall assume. If we set as the proper mass density

$$\rho_0=mn_0, \quad (19)$$

(14) becomes

$$\left(\rho_0\frac{dx^\mu}{d\sigma}\frac{dx^\alpha}{d\sigma}\right)_{,\alpha}=-\rho_0\Delta^\mu{}_{\alpha\beta}\frac{dx^\alpha}{d\sigma}\frac{dx^\beta}{d\sigma}. \quad (20)$$

Taking as the energy-momentum density tensor

$$T'^{\mu\nu}=\rho_0\frac{dx^\mu}{d\sigma}\frac{dx^\nu}{d\sigma}, \quad (21)$$

where the prime has the same significance as in I (20), one readily verifies that I (5) and hence I (21) are satisfied. In this way we see that the equations of motion and the field equations are consistent with one another.

§4

The preceding calculations show that the standpoint of flat space at any rate does not lead to internal contradictions. Let us now compare the theory based on this standpoint with the Einstein general relativity theory.

First of all, insofar as the three crucial tests[5] are concerned the two theories lead to the same predictions. Since the equations of motion have been taken the same, except for the independent variables, it is clear that relations among the coordinates, giving the motion of a particle or the path of a light ray, will also be the same in both theories. But more than that, the gravitational red shift will be the same in the two theories, although this is not quite as obvious.

The gravitational red shift in the general relativity theory is usually derived from the assumption that the interval ds connected with the period of oscillation of an atom dt remains unchanged if the atom is put into a gravitational field, together with the relation

$$ds^2=g_{44}dt^2$$

to which I (1) reduces if the space coordinates of the atom are fixed. It is clear, however, that this change in the period of oscillation must also be a consequence of the equations of motion of the particles constituting the atom. These equations being the same in the two theories, the predicted red shift will also be the same. (It should be mentioned that the equations of motion referred to will involve the electromagnetic field and may be of a quantum-theoretical nature.)

One way of describing the situation is to say that all the formal relations involving ds are taken over into the present theory but ds is not regarded as the geometrical interval. Putting the matter in this way makes the comparison of the two theories almost trivial. Thus one sees, for example, that the principle of equivalence is valid here as long as we deal with coordinates and not with intervals.

Indeed it appears at first sight as if the differences between the theories were only questions of conventions.

For example, if a small measuring rod is moved from one place to another in a gravitational field, the interval ds between its end points will remain constant, but the corresponding $d\sigma$ may change.[6] One may say that the length of the rod does or does not change, depending on the point of view. As far as the changes in the coordinate-differences of the end points are concerned, however, no disagreement exists. Incidentally, there is some analogy here to the change in length of a measuring rod with temperature. If all substances had the same temperature coefficients of expansion, one could assume that the length of the measuring rod is independent of the temperature and no contradictions would arise. This corresponds to the point of view of the Einstein theory.

[5] A. S. Eddington, *The Mathematical Theory of Relativity* (Cambridge, 1924).

[6] Reference 3, p. 818.

A similar example involving time is to be found in the case of the gravitational red shift.

Actually it turns out that there are other, more important, differences between the two theories. They are based upon the fact that, after all, the fundamental ideas of the two theories are different. *The present theory involves less relativity.* This may be seen from the following example:

Let us take for the moment the flat-space point of view. Suppose that in a given frame of reference, we have a static gravitational field described by $g_{\mu\nu}$. Then the velocity of light will, in general, be different from that in field-free space. The space containing the field can therefore be looked upon as a medium with an index of refraction different from unity. From the standpoint of special relativity one can expect that such a medium will exert a "drag" upon light, as in the case of the well-known experiment of Fizeau. If we go over to a second frame of reference moving uniformly with respect to the first (Lorentz transformation), it will be possible, by measurements carried out on light in this system, to determine its motion relative to the other system. Thus the static gravitational field determines an "absolute" frame of reference. This will be true whenever the velocity of light differs from its special-relativity value, even in the limiting case when the components $g_{\mu\nu}$ are constant in a part of space, so that there is no gravitational force (but the gravitational potential differs from zero).

Such a situation may account for the principal results of the "ether-drift" experiments of D. C. Miller.[7] A discussion of this question will be presented in a separate paper, since it involves some considerations which would lead us too far afield here.

If we consider the same situation from the point of view of the Einstein theory, we arrive at entirely different conclusions.[8] We cannot carry out a Lorentz transformation in the presence of a gravitational field but must first get rid of the latter by going over to a suitably accelerated frame of reference. Having obtained a coordinate system in which the components $g_{\mu\nu}$ are constant (at least for a small portion of space), we can always make a further coordinate transformation so as to give them their special relativity values. But then no light-drag or "ether-drift" can be observed.

§5

The question of viewpoint being discussed here has a bearing on the application of relativity theory to cosmology. Thus, if one takes seriously the interpretation in terms of flat space, one must give up the λ-term in the cosmological field equations,[9] since this prevents the $g_{\mu\nu}$ from going over into the $\gamma_{\mu\nu}$ at large distances from matter. Of course, any solution of the equations corresponding to an "expanding universe" must be interpreted as representing the expansion of a distribution of matter and radiation in a flat space. There appears to be no necessity for imposing the condition of homogeneity, as is usually done.

§6

In the Einstein general relativity theory gravitation is explained in terms of geometry. In the theory suggested here based on flat space, this geometrization of gravitation has been given up. Perhaps this may be regarded by some as a step backward. It should be noted, however, that the geometrization referred to has never been extended satisfactorily to other branches of physics, so that gravitation is treated differently from other phenomena. It is therefore not unreasonable to wonder whether it may not be better to give up the geometrical approach to gravitation for the sake of obtaining a more uniform treatment for all the various fields of force that are to be found in nature.

It might be remarked that the theory, as presented here, was obtained by a reinterpretation of the Einstein general relativity theory. This amounts to using the formalism of the latter as a means for obtaining a consistent set of equations. Such a procedure is not entirely satisfactory. It ought to be possible to build up the theory independently from suitable postulates. This will be attempted in the near future.

[7] D. C. Miller, Rev. Mod. Phys. **5**, 203 (1933).

[8] I am indebted to Professor Einstein for a helpful discussion of this question.

[9] A. Einstein, Berl. Ber. p. 142 (1917). As a matter of fact the existence of the nebular red-shift weakens the chief argument advanced there for introducing the λ-term.

Note on Ether-Drift Experiments

N. ROSEN
Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received October 30, 1939)

It is shown that it is possible to account for the principal results obtained by Miller in his "ether-drift" experiments, on the basis of a flat-space theory of gravitation, as due to the motion of the earth in the gravitational field of all the matter and energy in the universe.

SOME time ago, D. C. Miller presented a report[1] summarizing the large number of "ether-drift" experiments which he had carried out to determine the absolute motion of the earth through space. He concluded from his experiments that the earth is moving through space with a velocity of several hundred km/sec., but that the effect of this motion on the velocity of light (as measured on the earth) corresponds to a reduced velocity of about 10 km/sec., so that there is a drag-coefficient which he finds to have a value of about 0.051.

Unfortunately these results have not received satisfactory confirmation from other observers. Furthermore, Miller finds that the projection of the earth's translatory motion on the surface of the earth in the course of the day takes on directions which are unsymmetrical with respect to the north and south line. This seems difficult to explain on the basis of any simple theory. It may conceivably be an indication of some extraneous effect which accounts for his other results as well. However, the experiments appear to have been carried out very painstakingly and the results he obtained on the whole are remarkably consistent. Undoubtedly his experiments deserve serious consideration. Yet thus far no satisfactory explanation has been proposed.

In a recent paper by the writer,[2] the idea was considered of interpreting the formalism of the general theory of relativity in terms of flat space (special relativity). There, mention was made of the possibility of explaining the main result obtained by Miller on the basis of this idea. In the present note this question will be treated in some detail.

Let us suppose that, in a frame of reference S, we have a weak static gravitational field described by the gravitational potential $\varphi(\varphi \leqslant 0)$. Let us take a Galilean coordinate system in S with coordinates (x,y,z,t) so that the Euclidean line element $d\sigma$ is given by

$$d\sigma^2 = -dx^2 - dy^2 - dz^2 + dt^2, \tag{1}$$

the velocity of light in field-free space being taken as unity. Then the Riemannian line element of general relativity is given in first approximation by[3]

$$ds^2 = -(1-2\varphi)(dx^2+dy^2+dz^2)+(1+2\varphi)dt^2. \tag{2}$$

The velocity of light u is determined by the condition

$$ds = 0 \tag{3}$$

and one finds that it is independent of direction and is given in first approximation by

$$u = 1+2\varphi. \tag{4}$$

We see that the space can be looked upon as a medium having an index of refraction n, where

$$n = (1+2\varphi)^{-1} \doteq 1-2\varphi. \tag{5}$$

From the law of addition of velocities in the special theory of relativity it follows[4] that a medium of index of refraction n has a drag coefficient

$$k = 1-1/n^2. \tag{6}$$

This means in our case that, if there is an observer in a frame of reference S' moving with a uniform velocity v with respect to S, he will find that the velocity of light in the direction of motion will be

$$u_1' = u - kv \tag{7}$$

and that in the opposite direction

$$u_2' = u + kv, \tag{8}$$

[1] D. C. Miller, Rev. Mod. Phys. **5**, 203 (1933).
[2] N. Rosen, Phys. Rev. **57**, 149 (1940).
[3] A. Einstein, Berl. Ber. p. 688 (1916).
[4] R. Becker, *Theorie der Elektrizität*, Vol. II, p. 280.

where, in first approximation,

$$k=-4\varphi. \tag{9}$$

One can verify (7), (8), and (9), neglecting terms of the order of v^2, by carrying out a Lorentz transformation from coordinates (x,y,z,t) in S to (x',y',z',t') in S'. This leaves (1) unchanged in form when expressed in terms of the primed coordinates. From (2) and (3) then follow the preceding equations.

Thus we see that a static gravitational field provides a frame of reference with respect to which the uniform motion of an observer can be determined by experiments on light performed in his own system. It should be emphasized that this does not require the presence of a gravitational force; all that is necessary is that the potential φ be different from zero.

From the value 0.051 for k, as obtained by Miller, one has

$$\varphi=-0.013 \tag{10}$$

in the dimensionless units of general relativity theory. The question naturally arises: where does a potential of so large a magnitude come from?

The obvious answer seems to be that this potential is due to the distribution of matter throughout the universe. Since we do not have enough information about the latter to carry out an accurate calculation, we shall restrict ourselves to a rough estimate of the order of magnitude.

The gravitational potential at the center of a sphere of density ρ and radius R is given by

$$\varphi=-2\pi\kappa\rho R^2/c^2, \tag{11}$$

where κ is the gravitational constant and c the velocity of light in field-free space. Let us take for ρ the order of magnitude of the estimated mean density of matter in space:[5]

$$\rho=10^{-30}\text{g/cm}^3. \tag{12}$$

From (11) one can calculate the value of R which will give the value (10) for φ. It is found to be

$$R\sim 6\cdot 10^9 \text{ light-years}. \tag{13}$$

To obtain the same value of φ for a point not at the center of the sphere, one must take a larger value for R, of course. The value given in (13) is not unreasonable as an "effective radius" of the universe.

Thus we see that it is possible to account for Miller's observation of an "ether-drift," together with the value he finds for the drag-coefficient, on the basis of a flat-space theory of gravitation, as being due to the motion of the earth through the gravitational field of all the matter (and energy) in the universe. This field, the resultant of a large number of small contributions, is nearly static and serves to single out a frame of reference for determining "absolute" motion. The asymmetry referred to above remains unexplained.

It seems highly desirable to have more experimental work done on this question.

[5] E. Hubble, Astrophys. J. **79**, 8 (1934).

LETTERS TO THE EDITOR

Prompt publication of brief reports of important discoveries in physics may be secured by addressing them to this department. Closing dates for this department are, for the first issue of the month, the eighteenth of the preceding month, for the second issue, the third of the month. Because of the late closing dates for the section no proof can be shown to authors. The Board of Editors does not hold itself responsible for the opinions expressed by the correspondents.

Communications should not in general exceed 600 words in length.

The Propagation of Electron Waves in Ionic Single Crystals

The interaction between conducting electrons and the lattice structure of insulating crystals can be studied by the effects preceding and accompanying electric breakdown.[1] Especially striking is the dependence of the breakdown direction on the crystallographic orientation of the sample. A plate of rocksalt, for instance, cut parallel to the cleavage planes (100) and exposed to an intense homogeneous field perpendicular to it, does not break down first in the direction [100] of highest stress but shows a breakdown path parallel to the face diagonal [110] or, if overvoltage is applied, in the [111] direction. This "direction effect" can be explained by the potential structure of the crystal (Fig. 1): A surplus electron, in traveling through the lattice, encounters the lowest potential humps in the [110] direction and the smallest number of lattice points in the [111] direction, while the cube edge [100] is unfavorable in both respects. It might be expected that this directional effect is paralleled by a "magnitude effect." If the breakdown field strength for a plate cut parallel to (100) is $E_{\max}$, it should be $E_{\max}/2^{\frac{1}{2}}$ for a plate orientated parallel to the dodecahedral plane (110) because the face diagonal now points into the field direction. Accordingly a plate cut parallel to the octahedral plane (111) should break down at $E_{\max}\cdot 3^{\frac{1}{2}}/2$ if the breakdown field strength in [110] direction is decisive, that is,

$$\frac{E_{\max}[111]}{E_{\max}[110]}=(\tfrac{3}{2})^{\frac{1}{2}}.$$

Preliminary experiments carried out some time ago[2] did not fully confirm this expectation. While (111) plates showed the anticipated lower value of dielectric strength, the (110) plates withstood about the same stress as crystal plates cleaved in [100] direction. Both results were of dubious value because of the difficulty of faultlessly preparing very thin and polished plates of NaCl in directions differing from the natural cleavage plane.

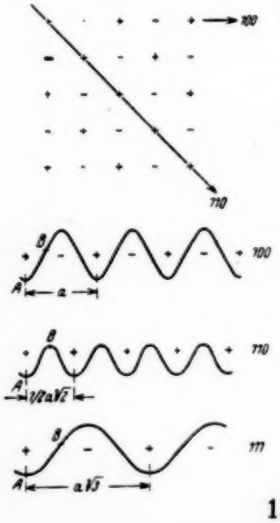

FIG. 1. Potential structure of the crystal.

In the meantime the question as to the existence of the magnitude effect has become of new importance. Following a note of Zener,[3] a wave-mechanical theory of the breakdown process has been carried through by Franz[4] based on the assumption that at very high field strengths the distinction disappears between conductor and insulator. According to the simple zone theory, the insulator is characterized by a completely filled lower band and an empty higher zone separated from the first one by about three volts. The forbidden interspace is created by Bragg reflections of the electron waves on the lattice planes, resulting in standing waves and thus preventing the acceleration of electrons. This mechanism becomes inefficient under very high field strengths because the energy difference between the filled band and the conducting levels can be overcome by accelerating the electrons through relatively few lattice planes, which are not sufficient for setting up strong reflected waves. In consequence, electrons leak through; the insulator becomes a conductor. If the number of planes traversed in the different lattice directions per unit length is the decisive factor for the intensity of reflection preventing conduction, a magnitude effect should be expected:

$$\frac{E_{\max}[111]}{E_{\max}[110]}=(\tfrac{3}{2})^{\frac{1}{2}}$$

as above, but furthermore

$$\frac{E_{\max}[111]}{E_{\max}[100]}=3^{\frac{1}{2}}$$

in contradiction to the direction effect.

Since every theory of electric breakdown has to know how far directional effects enter into the value of the field strength measured, the authors undertook a careful study of the existence of the magnitude effect in NaCl. The crystal plates were prepared by a new and simple method allowing one to secure faultless samples for every crystallographic direction in shortest time. After cutting thick samples in the right orientation out of synthetic single crystals, the material was ground down on glass plates under water to about 0.1 mm thickness. Wiped off with acetone, the crystal got a high polish, and heating in an electric furnace finished the treatment. Crystal plates cut parallel to the octahedral plane (111) showed some

tendency to crack, resulting in lower breakdown values. But after eliminating every disturbing effect, the measurements demonstrated without a doubt that *there does not exist a magnitude effect in rocksalt, at least not at room temperature.* In every direction tested, the breakdown strength is 1.65 million volts/cm with an error of about ±5 percent. This error is not so much due to the measuring technique, which would permit about two percent, as to inhomogeneousness of the material and roughness of the electrodes, resulting in locally higher stresses. Very slow raising of the voltage seems to have an equalizing effect, as already pointed out elsewhere;[5] in consequence the average breakdown field strength here reported lies about 10 percent higher than our earlier value.

The nonexistence of the magnitude effect disproves the wave-mechanical theory as given above. It is at all events hard to believe that the electrons of the Cl^- and Na^+ are not locally bound but move freely inside of their filled band through the whole crystal. But if no magnitude effect exists, the ionization theory of breakdown proposed by one of us and calculated in different ways by H. Fröhlich,[6] R. J. Seeger and E. Teller,[7] and W. Franz[4] also seems to come into difficulties. Our result shows that apparently the electrons do not feel the difference of the lattice directions until they have been accelerated to some extent. But this has to be expected, as Dr. Teller kindly advises us, because of the wave-length of the electrons. The breakdown strength of 1.65 million volts/cm corresponds to about 4.7×10^{-2} volt potential difference between the ions in [100] direction. The electrons start therefore with energies of the order of 10^{-1} volt corresponding to a wave-length of about 4×10^{-7} cm, that is, one electron covers about 15 ions in the [100] direction. This spread of the electron waves over a large lattice area makes them disregard the structure until they have shrunk by acceleration to the appropriate size.

A. VON HIPPEL
J. W. DAVISSON

Electrical Engineering Department,
Massachusetts Institute of Technology,
Cambridge, Massachusetts,
December 18, 1939.

[1] A. von Hippel, J. App. Phys. **8**, 815 (1937).
[2] A. von Hippel, Zeits. f. Physik **75**, 161 (1932).
[3] C. Zener, Proc. Roy. Soc. London **145**, 523 (1934).
[4] W. Franz, Zeits. f. Physik **113**, 607 (1939).
[5] R. C. Buehl and A. von Hippel, Phys. Rev. **56**, 941 (1939).
[6] H. Fröhlich, Proc. Roy. Soc. London **160**, 230 (1937).
[7] A. J. Seeger and E. Teller, Phys. Rev. **54**, 515 (1938).

The Nonexistence of Transuranic Elements

It is not clear why no atoms of atomic number greater than 92 are found in nature. The properties of the heaviest known nuclei do not suggest that all such atoms would be extremely unstable so that some reasonably long-lived alpha-emitting transuranic atoms might well be expected. Further, the absence of U^{236} is puzzling, for one would expect this nucleus to be as stable as U^{235} and U^{238}. It is striking that the periodic table of the elements ends with the only one that gives nuclear fission with thermal neutrons. This is possibly more than a coincidence and may be the clue to the explanation of the missing atoms.

According to the theory of Bohr and Wheeler[1] the probability of fission of any nucleus increases rapidly with increase of the energy of the captured neutron above a critical lower limit. The nuclei that give fission with slow neutrons are thus the ones to give the largest yields for bombardment by faster neutrons. Bohr and Wheeler calculate that of the reasonably stable nuclei (half-lives $>10^4$ yr.) only two should give fission with thermal neutrons. These are U^{234} and U^{235}. Pa^{231} should be so disintegrated by neutrons of energy greater than about 0.1 Mev. Atoms of lower atomic number have much higher critical energies so that the stable nuclei from Bi on down would require neutrons of extremely great energy for fission to be possible. Proceeding in the opposite direction to higher atomic numbers one finds that all reasonably expectable transuranic atoms should give fission with thermal neutrons. Further, U^{237} should also do this and U^{236} should require neutrons of only 0.1–0.2 Mev of energy. Even if these U nuclei and the transuranic nuclei had once been abundant they could have been destroyed by neutrons in a later phase of the development of the solar system. The presence of such neutrons is entirely hypothetical. Von Weiszaecker[2] has concluded that it is unlikely that there is any considerable concentration of neutrons in stellar atmospheres but that the neutron density must have once been high if the relative proportions of the different atomic species are those of an equilibrium at a high temperature.

The existence of the long-lived U^{225} (AcU) atoms seems at first glance to be in disagreement with this hypothesis. The postulated neutron bombardment, however, would have produced some U^{239} by capture. This is known to decay with a half-life of 23 min. by the emission of beta-rays. The resulting $_{93}EkaRe^{239}$ is most probably an ancestor of U^{235}, although the details of the intermediate transitions are not yet known. If the half-lives of the intermediate nuclei are short enough so that no important number of them would have been lost by fission some amount of U^{235} would have been maintained in a balance between this production from U^{238} and loss by fission.

Neutrons would also have been captured by Th^{232}. The Th^{233} so produced has been found to emit beta-rays and decay with a half-life of 25 min. The half-life of the resulting Pa^{233} is not known. The absence in nature of all of the atoms of the $4n+1$ family to which these belong is understandable if the presumably long-lived U^{237} was lost by fission without the replenishment as with U^{235}, and if neither the Pa^{233} nor any one of its descendants are long-lived. If the Pa^{233} decays by beta-emission to U^{233} this would have disappeared either because of being naturally short-lived or because of being especially susceptible to loss by fission.

These considerations will be developed in greater detail in a paper to be submitted to *The Physical Review* for publication.

LOUIS A. TURNER

Palmer Physical Laboratory,
Princeton University,
Princeton, New Jersey,
December 31, 1939.

[1] N. Bohr and J. A. Wheeler, Phys. Rev. **56**, 426 and 1065 (1939).
[2] C. F. von Weiszaecker, Physik. Zeits. **39**, 633 (1938).

A Measurement of Mesotron Lifetime

To obtain a proof of the mesotron decay hypothesis and to determine their lifetime without the necessity of making assumptions regarding their energy spectrum and the height of the layer where they are produced, the following observations have been made. The absorption curves of the cosmic radiation of vertical incidence were measured in lead from 0 to 15 inches Pb first on top of Mt. Mitchell, North Carolina (elevation 2040 meters, mean barometric pressure 60 cm Hg) and then at Durham, North Carolina (elevation 123.5 meters, mean barometric pressure 75 cm Hg). In the first series (Mt. Mitchell) the air column of 15 cm Hg = 203 g/cm² was compensated by a layer of graphite of 200 g/cm² filling the whole sensitive cone of the counter telescope.

In these measurements we have counted coincidences between four G-M counters. The slabs of absorbing lead were three times the width of the counters and about three times the sensitive length of the counters.

Some of the observed counting rates per hr. were thus:

Inches Pb	0	3	8	15
Mt. Mitchell	46.8	38.2	30.1	25.2
Durham	32.7	28.0	24.3	22.2

We estimate the statistical probable errors of the Mt. Mitchell data to be of the order of 1.2 percent and that of the Durham data to be of the order of 1.5 percent.

A similar series of observations at Mt. Mitchell with 10 cm of lead at the sides of one counter shows that the background showers were quite negligible in their effect on the course of the absorption curve. Furthermore, the number of such sidewise showers was small compared to the difference in counting rates in the two series of data. This large difference between the two series of measurements at Mt. Mitchell and Durham is therefore real and removes any doubt as to the reality of the difference in the absorption of the hard component in air as compared to dense matter. These measurements give, then, strong support to the mesotron decay hypothesis.

For the evaluation of the lifetime we use the following formula for the attenuation factor (reduction of intensity at the lower level):

$$A=(E/(E+Bx))\exp\left[\mu c^2/c\tau B\right],$$

where E is the energy at the lower level, B the specific energy loss in air (2×10^6 ev/(g/cm²) = 2.4×10^8 ev/km air), x the distance traveled, μ the rest mass of the mesotron, and τ the lifetime at rest. This formula takes into account the change of energy, i.e., the change in the Lorentz time factor along the path.

From the data of the table the lifetime comes out to be $\tau(\mu c^2/10^8\text{ ev})\times10^{-6}$ sec., where $\tau=1.2$, when A is calculated from the values at 8 and 15 inches Pb (corresponding to an average mesotron energy $E\sim4.4\times10^8$ ev) and $\tau=2.4$ for the values between 3 and 8 inches Pb (average energy $\sim2.8\times10^8$ ev). The latter value might already be falsified by a contribution from the soft component. The accuracy of the measurements is not yet sufficient to make certain the low value of 1.2 for τ, but it seems unlikely that it could be larger than 2. There would be no objection against a lifetime as small as 1.2×10^{-6} sec. from other observations (inclination measurements, barometer effect), if the production of the mesotrons takes place at a somewhat lower altitude than that of the maximum of the soft component or if the mean energy of the mesotrons is somewhat larger than heretofore assumed.

The value of the mesotron lifetime is then of the same order of magnitude as found in previous determinations. The actual measurements themselves are in reasonable agreement with the recent measurements of Rossi and his collaborators[1] insofar as a comparison can be made. From the fraction 26/34 of the mesotrons capable of traversing 12.5 cm or more of lead and which survive after a distance of 1900 meters one calculates a mean path of 7×10^5 cm as compared to Rossi's value of 8.5×10^5 cm.

The observations are being continued, and a full report will be published later.

W. M. Nielsen
C. M. Ryerson
L. W. Nordheim

Duke University,
Durham, North Carolina,

K. Z. Morgan

Lenoir Rhyne College,
Hickory, North Carolina,
December 19, 1939.

[1] B. Rossi, H. van Norman Hilberry and J. Barton Hoag, Phys. Rev. 56, 837 (1939).

Alpha-Particle Groups from the Disintegration of Beryllium by Deuterons

We have been investigating the energies of the alpha-particles from the reaction $Be^9(d,\alpha)Li^7$, with the variable pressure absorption cell described by N. M. Smith, Jr.[1] Since the mean range of these alpha-particles is about 3 cm in air at the deuteron energies available to us, and the effective length of the absorption cell is 6.699 cm, the operating pressures ranged from 15 to 35 cm.

Thin targets containing beryllium were prepared by volatilizing beryllium fluoride from a platinum wire held in the cooler parts of an oxygen-gas flame, and then inserting polished nickel target disks in the flame for 2 to 3 seconds. The cell was calibrated with a thin polonium source, which could be interchanged with the bombarded targets.

Figure 1 shows results with polonium, beryllium, and with a thin carbon target made by smoking one of the nickel disks in a gas flame. The scaling circuit was heavily biased so that counts were not recorded unless the alpha-particles passed through the 3-mm chamber at a velocity corresponding to their highest ionizing efficiency. The abscissae are pressures in the cell in centimeters of mercury. The ordinates for the carbon and beryllium curves are counts per 10^{-3} coulomb. The counting was automatically stopped by an integrator which tripped at 796 microcoulombs. The peak of the polonium curve corresponds to the counting of about 150 particles. The cell temperature for the polonium was 31.6°C, for the beryllium, 29°C.

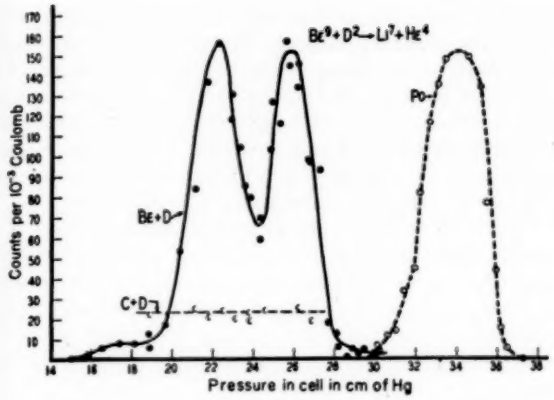

FIG. 1. Counts from Po, Be and thin carbon targets.

Table I shows the numerical results. Trial 2 of the table is that illustrated in Fig. 1. The energy release in this reaction may be calculated from the set of masses previously established in this laboratory,[1] and this calculated value is found to be 7.07 Mev, in excellent agreement with that obtained from the more energetic alpha-particle group. It is seen that there are two distinct energy groups, 0.31 Mev apart, and of about equal intensity at this deuteron energy. The points on the line marked C+D were taken with a carbon target, because the alpha-particles from the reaction $C^{13}(d,\alpha)B^{11}$ might appear in this region if our targets, in spite of our precautions, were contaminated with carbon.[2] This reaction, although observed by Cockcroft and Lewis[3] with 0.55-Mev deuterons apparently has too small a cross section to appear under our conditions. Visual observation of the oscillograph during the carbon counts showed that most of them were spurious. A similar negative result has been obtained from a target of potassium fluoride.

TABLE I. *Results. Energies in Mev.*

TRIAL	DEUTERON ENERGY	ENERGY OF SLOWER ALPHAS	ENERGY OF FASTER ALPHAS	Q	ΔE
1	0.351	4.34	4.64	7.05	0.30
2	0.326	4.35	4.67	7.10	0.32

These results indicate an excited state of Li^7 at 0.48 Mev, which is probably the same as that appearing in the proton groups from disintegration of Li^6 by deuterons.

This work was supported in part by a grant from the American Philosophical Society.

SAMUEL K. ALLISON
ELIZABETH R. GRAVES
LESTER S. SKAGGS

Ryerson Physical Laboratory,
University of Chicago,
Chicago, Illinois,
January 2, 1940.

[1] N. M. Smith, Jr., Phys. Rev. **56**, 548 (1939).
[2] G. T. Hatch, Phys. Rev. **54**, 165 (1938).
[3] J. D. Cockcroft and W. B. Lewis, Proc. Roy. Soc. **154**, 261 (1936).
[4] Reviewed by W. Maurer and J. B. Fisk, Zeits. f. Physik **112**, 436 (1939).

Internal Conversion in 6.4-Hour Cd

Valley and McCreary[1] have recently measured the internal conversion electron spectrum of 6.4-hour Cd. They point out that the ratio of the number of L_1 to the number of K conversions is approximately 1 for the 92-kev gamma-ray, while the conversion of a gamma-ray of the same energy in Ga^{67} gives the usual ratio of one to eight. The purpose here is to show that this result does not appear to be inconsistent with the theory.

The ratio of one to eight is predicted theoretically for gamma-radiation of energy much greater than the ionization potential of the electron.[2] In the case of Cd, the ionization potential of the K shell is about thirty percent of the gamma-ray energy (92 kev), therefore the above approximation is no longer valid. The effect is to reduce the probability of converting a K electron, as will be shown immediately. On the other hand, the L ionization potential is only three percent of the gamma-ray energy, so the L conversion is not reduced. The ratio of L to K conversions is therefore increased. Since the K-ionization potential of Ga is three times smaller than that for Cd, the reduction in the K-conversion probability for Ga should be much smaller.

The effect of a large K-ionization potential (I_K) is twofold:

(1) It reduces the kinetic energy ($\epsilon_K = h\nu - I_K$) of the ejected K electron and therefore reduces the density of energy states into which it can go, since the density is proportional to the square root of the kinetic energy. This reduction in the statistical weight of the transition decreases the transition probability.

(2) It has been shown by Taylor and Mott[3] that for long wave-length gamma-radiation, internal conversion is almost entirely due to direct interaction of the K (or L) electrons with the nucleus. If the radiation is assumed to be dipole, the potential acting on the electron is of the form P/r^2, where P is the effective dipole moment of the nucleus. Therefore the matrix element for the electronic transition is $P\int \psi_0(1/r^2)\psi r^2 dr = P\int \psi_0\psi dr$. For dipole radiation, the final electronic wave function, ψ, is a P function, which has a zero of order r/λ, $\lambda = \hbar/(2m\epsilon)^{\frac{1}{2}}$, at the nucleus. The matrix element is therefore proportional to $(1/\lambda)\int \psi_0\varphi r dr$, and, since ψ_0 and $\varphi(=\lambda\psi/r)$ are slowly varying functions of r that simply determine the effective limits of integration, this is proportional to $(1/\lambda)\int_0^x r dr = x^2/2\lambda$. The upper limit, x, is equal to the smaller of a_0 and $\lambda/2$, where a_0 is the radius of the K or L shell as the case may be. The transition probability is proportional to the square of this matrix element, i.e., proportional to x^4/λ^2.

The ratio of L to K conversion is given by

$$w_L/w_K = (\tfrac{1}{8})x_L^4\lambda_K^2/x_K^4\lambda_L^2(\epsilon_L/\epsilon_K)^{\frac{1}{2}} = (\tfrac{1}{8})(x_L/x_K)^4(\lambda_K/\lambda_L)^3,$$

where the factor $\frac{1}{8}$ is the value of the probability for $\lambda_L = \lambda_K$ (i.e., for $h\nu \gg I_K$). For Cd,[1] $a_K = \hbar^2/Zme^2 \approx 1.1$, $\lambda_K \approx 4.7$, $a_L = 4a_K \approx 4.4$, $\lambda_L \approx 4$ in units of 10^{-10} cm. Since $a_K < \lambda_K/2$ and $a_L > \lambda_L/2$, the ratio becomes

$$(w_L/w_K)_{Cd} = (1/128)(\lambda_K^3\lambda_L/a_K^4) \approx 2.$$

On the other hand, for Ga, $a_K \approx 1.7$, $\lambda_K \approx 4.2$, $a_L \approx 6.8$, $\lambda_L \approx 4$ in the same units, and $(w_L/w_K)_{Ga} \approx 2/8$.

Since the estimates are extremely rough, it is not surprising that these values differ from the experimentally observed values[1] of 1 and $\frac{1}{3}$ by a factor of two. However, the ratio of the ratios, $(w_L/w_K)_{Ga}/(w_L/w_K)_{Cd}$, might be expected to be somewhat independent of the systematic errors in the rough estimate, and it turns out to be approximately $\frac{1}{3}$, in good agreement with the experimental value.

R. G. SACHS

The George Washington University,
Washington, D. C.,
December 21, 1939.

[1] G. E. Valley and R. L. McCreary, Phys. Rev. **55**, 322 (1939).
[2] H. R. Hulme, Proc. Roy. Soc. **A138**, 643 (1932).
[3] H. M. Taylor and N. F. Mott, Proc. Roy. Soc. **A142**, 215 (1933).

On the Neutron's Magnetic Moment

As announced previously,[1] we have calculated in detail the influence of magnetization on the transmission of neutrons through polycrystalline iron. Since the coherent scattering is concentrated in Debye-Scherrer rings and since the transmission effect depends in a sensitive manner on the scattering angle, the final result differs radically from that previously calculated[2] for randomly oriented atoms.

The only essential uncertainties in the theoretical evaluation occur from minute deviations from magnetic saturation and from our incomplete knowledge of the distribution of the magnetically active shell in the iron atom. In these calculations the neutron's magnetic moment was assumed to be two nuclear Bohr magnetons. The depolarization effects due to incomplete saturation can only diminish the transmission effect.

To determine the extension of the magnetically active shell we have used various density functions presented in the literature. Slater's[3] hydrogenic functions give so small a form factor for the magnetic scattering that no effect could be expected. Adopting the analytic representation[4] for the Hartree distribution of the 3d electrons in Cu, to Fe, we obtained values for the transmission effects which are very much smaller than any observed previously and differ from those given by Bloch and Alvarez[5] by a factor of approximately $\frac{1}{3}$. The discrepancy is considerably enlarged due to the presence of depolarization effects which manifested themselves experimentally in a deviation from the law: added transmission proportional to the square of the thickness.

Attempts to explain the discrepancy can in our opinion only be based upon the two following possibilities: (1) The form factor used in these calculations is too small because either the charge distribution is too extended or due to some kind of spin orbit coupling the distribution of magnetic moment is contracted with respect to the distribution of charge. (2) Since only the gyromagnetic ratio for the neutron has been determined experimentally,[5] it is still possible to assume that the neutron's spin as well as its moment may be three times[6] as large. In this case the transmission effect observed would agree with the calculated value if a reasonable correction for the depolarization effect is introduced.

We are conscious that a value for the neutron's spin larger than $\frac{1}{2}$ would seriously change all theoretical considerations concerning neutron-proton-scattering.[7] We shall enter into this question more fully in a subsequent paper in which the theoretical aspects of the new hypothesis will be discussed. Here it may suffice to say that present experimental evidence[8] is far from being completely in favor of the commonly presented theoretical view.

Observations on the paramagnetic scattering[9] of slow neutrons allow the determination of the neutron's *magnetic moment* directly with comparative case. Assuming a spin of $\frac{1}{2}$ and using the gyromagnetic ratio as given by Bloch and Alvarez,[5] the scattering cross section of free Mn^{++} for a very long neutron wave-length should be approximately 100×10^{-24} cm^2. Observations on differential cross sections under small angles would clearly decide whether the discrepancy here discussed is due to cause (1) or (2) mentioned above.

O. HALPERN
M. H. JOHNSON

New York University,
University Heights, New York,
November 14, 1939.

[1] O. Halpern and Th. Holstein, Phys. Rev. **55**, 601 (1939); Halpern, Hamermesh and Johnson, Phys. Rev. **55**, 1125 (1939).
[2] J. Schwinger, Phys. Rev. **51**, 544 (1937).
[3] J. C. Slater, Phys. Rev. **36**, 57 (1930).
[4] J. C. Slater, Phys. Rev. **42**, 33 (1932).
[5] F. Bloch and L. W. Alvarez, Phys. Rev. **57**, 110 (1940).
[6] To fit the spin of the deuteron and the demands of the Fermi statistics.
[7] J. Schwinger, Phys. Rev. **52**, 1250 (1937); H. H. Goldsmith and L. Motz, Phys. Rev. **53**, 947 (1938).
[8] M. Goldhaber, Nature **137**, 824 (1936); M. A. Tuve and L. R. Hafstad, Phys. Rev. **50**, 490 (1936); H. v. Halban, Nature **141**, 644 (1938); E. Amaldi *et al.*, Phys. Rev. **56**, 881 (1939); T. Goloborodko and A. Leipunski, Phys. Rev. **56**, 891 (1939).
[9] O. Halpern and M. H. Johnson, Phys. Rev. **55**, 898 (1939).

On the Magnetic Moments of Light Nuclei

That the discrepancies[1] between theoretical and experimental values of nuclear magnetic moments may be due to oversimplifying assumptions in the theory concerning the ground states of the various nuclei has often been pointed out. A detailed investigation of an added interaction of the form

$$V=J(r)\left(3\frac{\boldsymbol{\sigma}_1\cdot\mathbf{r}\;\boldsymbol{\sigma}_2\cdot\mathbf{r}}{r^2}-\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2\right)$$

in the case of the deuteron[2] led to consideration of this coupling as an explanation. Its possible effect on the magnetic moments of several light nuclei has been suggested[3] and the cases of Li^6 and N^{14} have recently been discussed.[4]

V has nonvanishing matrix elements between states having the same total angular momentum and the same spatial symmetry (i.e., characterized by the same "partition" quantum number). The sequence of observed moments for the stable nuclei having an odd number of protons and the same number of neutrons (H^2, Li^6, B^{10}, N^{14}) suggests an increasing proprtion of $^3D_1$ in the ground state, as provided by this coupling term, and calculations made on the basis of the Hartree model qualitatively confirm this suggestion. Furthermore, the $g=1.788$ and probable

$J=\frac{3}{2}$ values for B^{11} can be understood in the same way. Spherically symmetric forces alone lead to a ground state[5] of $J=\frac{1}{2}$. There are, however, two nearby states with $J=\frac{3}{2}$, and a rough calculation makes it seem reasonable that they interact sufficiently for one of them to become the ground state. In the language of LS coupling the ground level is then a mixture of $^2P_{3/2}$ and $^2D_{3/2}$, and the correct magnetic moment is obtained theoretically if there is nearly one-half as much ^2D as ^2P.

With the modification produced by the interaction V the Hartree model is then consistently correct in its predictions of spins for the light nuclei, giving the observed[6] $J=\frac{1}{2}$ for C^{13} in contrast to the α-particle model.[7] Quantitative agreement cannot be expected, and there is as a matter of fact no justification for detailed calculation with single particle wave functions, but the signs and rough magnitudes of the theoretical magnetic moments are in accord with observation. One serious discrepancy has not been so explained. The experimental value of $\mu=3.25$ for Li^7 is larger than the theoretical value as given by the Hartree or any other simple model. The angle dependent coupling has no appreciable effect, since there is only one low energy term with $J=\frac{3}{2}$. Until this case is resolved one cannot be too optimistic about understanding observed magnetic moments in terms of the interactions hitherto investigated in connection with the problem.

It is unfortunate that the Alvarez-Bloch result for the neutron moment is not sufficiently precise to constitute a criterion for the validity of these considerations. The coupling V, in accounting for the quadripole moment of the deuteron, depresses the value of μ_N, while their result gives no evidence of departure from pure additivity of proton and neutron moments in the deuteron.

M. PHILLIPS

Department of Physics,
Brooklyn College,
Brooklyn, New York,
December 1, 1939.

[1] M. E. Rose and H. Bethe, Phys. Rev. **51**, 205 (1937).
[2] J. Schwinger, Phys. Rev. **55**, 235 (1939); H. Bethe, Phys. Rev. **55**, 1261 (1939).
[3] Millman, Kusch and Rabi, Phys. Rev. **56**, 165 (1939).
[4] M. E. Rose, Phys. Rev. **56**, 1064 (1939).
[5] M. E. Rose and H. Bethe, Phys. Rev. **51**, 993 (1937).
[6] C. H. Townes, Phys. Rev. **56**, 850 (1939).
[7] R. G. Sachs, Phys. Rev. **55**, 825 (1939).

The Neutron-Deuteron Scattering Cross Section*

We have calculated the elastic-scattering cross section of neutrons by deuterons for energies so small (thermal neutrons) that only the s part of the incident wave is appreciably different from the corresponding part of a plane wave.

In the coordinate system of the center of mass of the three particles, defined by the transformation,

$$\mathbf{r}=\mathbf{r}_1-\tfrac{1}{2}(\mathbf{r}_2+\mathbf{r}_3),$$
$$\boldsymbol{\rho}=\mathbf{r}_2-\mathbf{r}_3,$$
$$\mathbf{R}=\tfrac{1}{3}(\mathbf{r}_1+\mathbf{r}_2+\mathbf{r}_3)$$

(1 refers to the incident neutron and 3 to the proton) the wave equation is

$$[-(\hbar^2/2M)(2\nabla_\rho^2+\tfrac{3}{2}\nabla_r^2)+V_{12}+V_{13}+V_{23}-E]\Psi_S(r,\rho)=0.$$

The subscript S is the spin quantum number, and $E=E_0+E'$, where E_0 is the deuteron binding energy and E' is the kinetic energy of the incident neutron in the center of mass system. For the interactions V_{ij} we have taken the general form[1]

$$V_{ij}=-[(1-g-g_1-g_2)P_{ij}+gP_{ij}Q_{ij}+g_1+g_2Q_{ij}]V_0\exp[-r_{ij}^2/a^2],$$

in which $Q_{ij}=\frac{1}{2}(1+\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j)$ is the spin exchange operator (σ_i, σ_j are the Pauli matrices of amplitude 1) and P_{ij} is the coordinate exchange operator. In the present calculation V_0 was taken equal to 72 mc^2 and a, the range equal to $0.25\hbar/(Mm)^{\frac{1}{2}}C$ cm. Because of the complexity of the problem we have neglected any polarization of the deuteron by the neutron.

The following expansions were used:

$$\Psi_S(r,\rho)=F_0(r)\psi_0(\rho)\chi_S(123)+\chi_S(123)\int_0^\infty F(r,E)\psi(\rho,E)dE,$$
$$P_{12}Q_{12}\Psi_S(r,\rho)=G_0(r)\psi_0(\rho)\chi_S(123)+\chi_S(123)\int_0^\infty G(r,E)\psi(\rho,E)dE.$$

$\psi_0(\rho)$ is the ground state deuteron wave function, $\chi_S(123)$ is the total spin wave function for a definite value S of the spin, and $\psi(\rho,E)$ is the excited deuteron wave function.

On substituting these expansions in the wave equation and integrating over the deuteron coordinates, one obtains two equations which can be combined by subtraction to give

$$-\frac{3}{4}\frac{\hbar^2}{M}\nabla_r^2\phi(r)+\sum_{\text{Spins}}\sum_{i=2,3}\int d\rho\chi_S\psi_0(\rho)V_{1i}\times\{\phi(r)\psi_0(\rho)\chi_S-(PQ)_{12}(\phi(r)\psi_0(\rho)\chi_S)\}=(E-E_0)\phi(r).$$

In this equation the approximation

$$\Psi_S(r,\rho)-(PQ)_{12}\Psi_S=[F_0(r)-G_0(r)]\psi_0(\rho)\chi_S-(PQ)_{12}[F_0(r)-G_0(r)]\psi_0(\rho)\chi_S\equiv\phi(r)\psi_0(\rho)\chi_S-(PQ)_{12}\phi(r)\psi_0(\rho)\chi_S$$

was used.

Since polarization has been neglected, $\phi(r)$ is correctly symmetrized to take account of exchange effects. The scattering cross section can therefore be found from that solution $\phi(r)$ which has the asymptotic form

$$\phi(r)\to\exp[i\mathbf{k}\cdot\mathbf{r}]+r^{-1}e^{ikr}f(\theta).$$

The total scattering cross section is then given by

$$\sigma_t=\int\sigma(\theta)d\omega=2\pi\int|f(\theta)|^2\sin\theta d\theta.$$

If we substitute the general expression for V_{1i} in the equation for $\phi(r)$, the sum over the spins can be easily performed, and the differential equation can be replaced by an integral equation with the correct asymptotic form.[2] This integral equation has to be solved numerically. The calculation is greatly simplified by the use of Gaussian potentials and a Gaussian function for $\psi_0(\rho)$. In the present calculation we chose

$$\psi_0(r)=[\pi^{\frac{1}{2}}a^3(1.23)]^{-\frac{1}{2}}(0.80e^{-0.6r^2/a^2}+0.12e^{-0.65r^2/a^2})$$

which gives the binding energy of the deuteron to within about 5 percent.

The presence of the Gaussian functions in the integrand insures a rapid enough convergence so that the integral equation can be replaced by a finite set of linear algebraic equations. It was found in this calculation that for $r \lesssim 13a$ the contributions to the integrals were practically zero.

Since both doublet (d) scattering, $S=\frac{1}{2}$ and quartet (q), $S=\frac{3}{2}$ scattering are present, we had to solve two sets of equations for the two values of the total spin. For thermal neutrons ($k=0$) and for the special choice of potential given by $g=0.2$, $g_1=g_2=0$,[1] we obtained

$$\phi_d(13a)=1.0546; \quad \phi_q(13a)=0.749$$

and

$$|f_d(\theta)|=|\phi_d(13a)-1|(13a)=1.59\times10^{-13} \text{ cm},$$
$$|f_q(\theta)|=|\phi_q(13a)-1|(13a)=7.302\times10^{-13} \text{ cm}.$$

From this the scattering cross section was found to be

$$\sigma_t=\frac{4\pi}{3}(2|f_q(\theta)|^2+|f(\theta)|^2)$$
$$=4.57\times10^{-24} \text{ cm}^2.$$

The recent experiments of Carroll and Dunning (to be published) give the value 5.7×10^{-24} for this scattering cross section. If one takes molecular binding into account, this value has to be reduced by a factor which lies between 1.5 and 2.[3] At best, then, the theoretical value is about 20 percent too large. This discrepancy may be accounted for by polarization effects.

We should like to express our thanks to Mr. Jerome Rothstein of Columbia University for aiding in the numerical work.

LLOYD MOTZ

Columbia University,
New York, New York,

JULIAN SCHWINGER**

The University of California,
Berkeley, California,
December 18, 1939.

* Publication assisted by the Ernest Kempton Adams Fund for Physical Research of Columbia University.
** National Research Fellow.
[1] G. Breit and E. Feenberg, Phys. Rev. **50**, 850 (1936).
[2] N. E. Mott and H. S. W. Massey, *The Theory of Atomic Collisions* (Oxford Univ. Press, New York, 1933), Chap. IV.
[3] L. I. Schiff, Phys. Rev. **52**, 149 (1937).

Absorption Spectrum of Single-Crystal ZnS Phosphors

A knowledge of the absorption spectrum of a phosphor is necessary if one is to understand the process by means of which radiation is emitted by the phosphor. In view of the experimental difficulties encountered in obtaining the absorption spectrum of a powder, it seems desirable to study the absorption spectrum of a ZnS phosphor in the form of a single crystal; especially so, since there seems to be some uncertainty as to the nature of the absorption process in activated ZnS.[1]

Several ZnS single crystals, approximately one millimeter square and a few tenths of a millimeter in thickness, were obtained through the kindness of Dr. R. P. Johnson and Mr. F. B. Quinlan of the General Electric Company. The crystals were made by firing in vacuum a precipitated

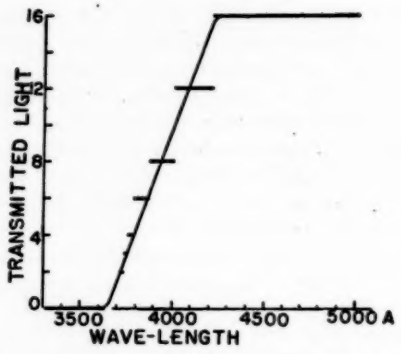

FIG. 1. Relative intensity of the light transmitted by a ZnS·Cu single-crystal phosphor *vs.* the wave-length, in angstrom units.

ZnS with about 0.004 percent Cu. Each of three crystals was mounted behind a small hole in black paper and then placed in front of the slit of a small Littrow quartz spectrograph. Light from a condensed iron spark was focused on the crystal by means of a quartz lens. The spectra were photographed on Eastman Spectroscopic Plates, Type IV–O. Three identical two-minute exposures were taken with the ZnS crystal in position. The crystal was then removed and six comparison spectra of 1, 2, 4, 8, 16, and 32 seconds exposure were placed on the plate so as to bracket each of the original spectra. The intensities of the lines in the original spectra were determined visually by matching densities with one of the comparison spectra, whose exposure time was then taken as a measure of the intensity in question.

The absorption curves obtained in this way for three of the ZnS single-crystal phosphors were quite similar; a typical curve, shown in Fig. 1, indicates that for this phosphor, absorption is partial for wave-lengths less than 4250A and is complete for wave-lengths less than 3650A. This absorption curve is in good agreement with the curve obtained by Kitchener[1] from a ZnS·Cu powder, using the diffuse reflection method. Gisolf has reported[2] a secondary absorption peak on the long wave-length side of the absorption edge of ZnS phosphors; nothing of this nature is clearly shown by the accompanying curve, although there seems to be a slight asymmetry in the absorption edge which may be due to an unresolved secondary absorption peak. Nor was any evidence found of absorption bands in the region 4200–5000A, such as have been reported by Kroger[3] for ZnS·MnS mixed phosphors. Any such partial absorption which would reduce the intensity of the transmitted light by as much as one-fourth should have been detected.

CHARLES J. BRASEFIELD

U. S. Coast Guard Academy,
New London, Connecticut,
December 22, 1939.

[1] See, for example, page 97 of "Luminescence, A General Discussion held by the Faraday Society, September 1938" (reprinted from the Transactions of the Faraday Society, Vol. 35); and references mentioned there.
[2] J. H. Gisolf, Physica **6**, 88 (1939).
[3] F. A. Kroger, Physica **6**, 374 (1939).